

**IMPROVED PROCEDURES FOR QUANTIFYING
KEY METEOROLOGICAL EFFECTS ON
AMBIENT OZONE DATA**

**Final Report
Contract No. A832-134**

Prepared for:

Research Division
California Air Resources Board
1102 Q Street
P.O. Box 2815
Sacramento, CA 95812

Submitted by:

Systems Applications International
101 Lucas Valley Road
San Rafael, California 94903

Prepared by:

Till E. Stoeckenius
Chris Daly
A. Belle Hudischewskyj
Art M. Noda

June 1991

Disclaimer

The statements and conclusions in this report are those of the contractor and not necessarily those of the California Air Resources Board. The mention of commercial products, their source or their use in connection with material reported herein is not to be construed as either an actual or implied endorsement of such products.

Executive Summary

Ozone concentrations in California's urban air basins exhibit considerable day-to-day and year-to-year variability. A large portion of this variability may be attributed to fluctuations in meteorological conditions. From the standpoint of air quality management, this intimate relationship between the weather and ozone has two principal consequences:

Fluctuations in the frequency and severity of weather conditions conducive to the formation of high ozone concentrations result in large differences in concentrations from one year to the next. These fluctuations effectively mask any underlying trend in ozone that may have resulted from changes in the geographical distribution and amount of precursor emissions.

Meteorological conditions may vary significantly from one high ozone episode to the next. Because of the expense of obtaining data for applying photochemical grid models, it is only feasible to study the effects of emission control strategies on the basis of data collected for a small number of episodes. Since the effectiveness of a particular control strategy in reducing ozone concentrations will vary depending on weather conditions, it is difficult to predict the overall effect of a control strategy on the basis of results from just a few modeled episodes.

In light of these consequences, a series of analyses were carried out with the following goals in mind:

Quantification of the effects of meteorological conditions on ozone concentrations.

Development of information to be used in the selection of ozone episodes for control strategy modeling and meteorological adjustment of ozone trends.

These methodologies will allow air quality managers to deal with the problems noted above by providing them with the information needed to (1) calculate year-to-year ozone trends that have been adjusted for meteorological fluctuations and thus more clearly indicate the effects of changes in precursor emission patterns and (2) make informed decisions on the representativeness of ozone episodes for purposes of emission control scenario development.

During the planning stages of this study, it was determined that the best strategy for achieving the goals outlined above would be to develop a methodology for grouping days into categories such that days within the same category will have similar meteorological and air quality characteristics and thus be representative of one another in the sense described above. Given a method of identifying such "source history" categories, one could then select a limited number of days from each category for the purpose of modeling air quality responses to alternative emission control scenarios. In this way, model results would be available for a set of days that represent the range of episode conditions affecting the air basin under study. Furthermore, meteorologically adjusted air quality trends could be computed by calculating the change in ozone concentrations from one year to the next on days falling within the same source history category. Since such days are associated with nearly identical meteorological conditions, any year-to-year changes in concentrations must be attributed to changes in precursor emission levels.

A pilot study was conducted to determine the feasibility of the approach. The pilot study focused on data collected in the South Coast Air Basin (SOCAB) for the period May–October, 1983–1985. To avoid difficulties imposed by changes in emission patterns resulting from weekend work and travel anomalies, only data for Tuesdays, Wednesdays and Thursdays were included in the analysis. Statistically robust ozone concentration measures were obtained by averaging the daily maximum ozone concentration over stations in each of nine basin subregions. Subregions were defined to include groups of nearby stations with highly correlated ozone concentrations. To provide a starting point in the search for coherent source history patterns, Mel Zeldin of the South Coast Air Quality Management District reviewed the spatial distribution of ozone and the meteorological characteristics of each day included in the pilot study period and identified eight potential source history categories:

Typical Pattern: Days in this pattern are characterized by westerly or north-westerly coastal winds that transport ozone and precursors through the San Gabriel Valley and into the eastern portions of the basin.

Eddy Pattern: Days in this pattern are characterized by coastal winds with a pronounced southerly component that produce a greater transport of ozone and precursors into the San Fernando Valley and upslope into the mountains than occurs on Typical Pattern days.

Southern Route Pattern: This pattern includes days with more northerly winds that tend to reduce the easterly flow through the San Gabriel Valley subregion and push ozone and precursors further into Orange County.

Offshore Pattern: This pattern includes days in which a weak offshore (i.e., negative) coastal to inland pressure gradient reduces the eastward extent and vigor of the sea breeze circulation, resulting in ozone levels that are higher in the coastal and adjacent metropolitan areas and lower in the easternmost portions of the basin than on Typical Pattern days.

Three additional categories representing combinations of the above patterns were also identified:

Partial Eddy Pattern: Days in this category exhibit spatial ozone distributions that are similar to those observed in the Eddy Pattern but with slightly lower values in the San Fernando Valley and in the mountains. However, the southerly component winds characteristic of the Eddy Pattern are not present on these days.

Typical Pattern with Eddy Winds: A few days were identified in which coastal winds had a definite southerly component but the distribution of relative ozone concentrations was similar to those found on Typical Pattern days.

Partial Southern Route: Days in this category exhibit some of the meteorological characteristics of Southern Route days, with a weak push of pollutants into northern Orange County but with sufficient afternoon westerly winds to produce the usual basin-wide maximum concentrations in the San Gabriel Valley that are characteristic of Typical Pattern days. This scenario results in higher relative concentrations in the vicinity of Pico Rivera than are observed on Typical Pattern days.

Finally, all days in which no subregion average daily maximum ozone concentration in excess of 8 pphm occurred were assigned to an eighth, low ozone, category.

The meteorological and air quality characteristics of the above categories were analyzed in detail in the remainder of the pilot study.

Results of the pilot study indicated that at least some of the source history categories initially identified by Zeldin appear to be associated with distinct and coherent spatial ozone concentration patterns and meteorological features. However, the distributions of concentrations and meteorological parameters exhibited a great deal of overlap from one category to the next, making it difficult to clearly identify the unique nature of each one.

One measure of the power of the source history categories to differentiate between types of high ozone days is the percentage of the variance in the daily maximum ozone concentration in each subregion that can be explained by the categories themselves or by the categories in combination with within-category regressions against meteorological parameters. Ignoring the low ozone days (i.e., those in the eighth category), our pilot study results indicate that the division of days into the remaining seven source history categories explains 39 to 47 percent of the variance in the coastal and adjacent metropolitan subregions, 22 to 52 percent of the variance in the subregions lying further inland, and just 6 percent of the variance in the San Fernando Valley. When regression models are used to further define the relationship between ozone concentrations and meteorological conditions within each

source history category, the percentage of variance that can be explained increases to a total of 64 to 77 percent in the coastal and metropolitan subregions, 74 to 80 percent in the inland subregions and 60 percent in the San Fernando Valley. By way of comparison, simple linear regressions using the 850 mb temperature without identification of source history categories are capable of explaining roughly 50 to 70 percent of the variance in the daily maximum ozone concentration averaged over several inland stations.

In the second part of our study, we undertook analyses designed to provide additional support for the pilot study conclusions by obtaining additional information about the distributions of meteorological parameters and ozone concentrations within each source history category and comparing these distributions between categories. Two aspects of the categories were examined:

Diurnal ozone concentration profiles normalized by the daily average concentration were calculated for each day at each subregion and averaged over days in each source history category.

Air parcel trajectories were examined for a set of days selected from each category. Back trajectories were calculated from the arrival of parcels at several monitors at the time of the afternoon ozone maximum to early on the morning of the same day.

Results of these analyses indicated that the Eddy and Southern Route categories were the most well defined of Zeldin's categories, that Partial Eddy and Partial Southern Route days are very similar to Eddy and Southern Route days, respectively, and that the remaining days exhibit a diverse range of flow patterns and ozone distributions.

Application of source history categories to episode selection and meteorological trend adjustment requires the development of meteorological criteria which can be used to classify each day into the appropriate category. As a result of the difficulties encountered during the pilot study in identifying sufficiently unique meteorological signatures for each category, we turned our efforts towards the establishment of meteorological criteria that would allow us to identify the two most distinct categories: Eddy and Southern Route. Rosenbaum (1990) performed extensive analyses of the meteorological conditions associated with these categories, including the examination of numerous surface and upper-air parameters not considered in our pilot study. Rosenbaum succeeded in finding a series of criteria which accurately identify days with spatial ozone patterns that match those associated with Zeldin's Eddy or Partial Eddy and Southern Route or Partial Southern route source history categories. Although most days meeting Rosenbaum's criteria had the expected spatial ozone patterns, only about half of all days with such ozone patterns were found to also meet Rosenbaum's criteria. The remaining days, along with days in the Typical, Typical with Eddy Winds and Offshore source history categories, were assigned to an "Uncertain" meteorological category. Although the presence of a

large group of days in the Uncertain category is undesirable for some applications, the restrictive nature of Rosenbaum's meteorological classification procedure is useful in the sense that it insures that nearly all of the days which do meet the criteria will exhibit the expected ozone pattern. We therefore adopted Rosenbaum's criteria for use in our study and used them to categorize all days included in the study period (Tuesday - Thursday, May - October, 1980 - 1988).

One would expect to see differences between Rosenbaum's Eddy and Southern Route meteorological categories in the relationship between meteorological conditions and ozone concentrations at various locations within the SOCAB. We performed two different regression analyses in an attempt to identify which meteorological variables are most closely correlated with ozone concentrations within each category. Although results of the two regression approaches are somewhat inconsistent, temperature variables were found to be closely related to ozone formation in all cases.

A potential use of the meteorological categorization procedure is in the adjustment of ozone concentration trends to account for variations resulting from changes in weather conditions. A preliminary evaluation of this potential was performed as follows: Day-to-day variations in ozone concentration for days within each category were accounted for using a series of linear regression analyses. A separate regression was performed for each meteorological category using a set of seven key meteorological variables. Separate regression equations were developed for each of seven overlapping three-year intervals (1980 - 1982, 1981 - 1983, ..., 1986 - 1988). Three-year periods were used to insure that each regression was based on a sufficient number of data points. The resulting regression equations were then used to calculate adjusted seasonal mean ozone concentrations for each 3-year period for each meteorological category by substituting into the equations the long-term (9-year) average values of the meteorological regressor variables. The resulting within-category adjusted concentrations can be interpreted as the values which would have been observed had meteorological conditions during each 3-year period for days in the category corresponded to climatological norms. In the final step of the adjustment process, adjusted seasonal mean concentrations over all meteorological categories were calculated by averaging the adjusted concentrations across categories, assuming that the frequency of occurrence of each category during each 3-year interval is equal to the long-term (9-year) average frequency of occurrence. In this way, the influence on seasonal mean ozone of year-to-year differences in the number of days falling in each category is accounted for and the resulting adjusted concentrations reflect not only the effects of meteorological variations within each category but also between categories.

Unfortunately, the within-category regression equations on which the meteorological adjustment calculations were based proved to fit the data quite poorly in many cases. As a result, the adjusted seasonal mean concentrations were unreliable and the adjusted trends were just as noisy if not more noisy than the unadjusted trends. In addition, much of the year-to-year variability in concentrations was found to be removed simply by the calculation of 3-year running averages, leaving only relatively

small variations for the adjustment procedure to eliminate.

We compared results from the preliminary category-based trend adjustment procedure described above to results obtained from an alternative adjustment procedure suggested by Davidson et al. (1985). This procedure makes use of the observation that SOCAB ozone concentrations are closely correlated with the morning 850 mb temperature. Since no meteorological categories are used, sufficient data were available to apply the adjustment method to individual years. Results obtained indicate that some of the year-to-year variations in seasonal mean ozone are accounted for by variations in temperature. However, considerable correlation between the adjusted and unadjusted trends remains, suggesting that additional meteorological variability remains unaccounted for. Running three-year averages of the temperature adjusted trends were calculated to provide results that could be compared directly to those obtained using the category-based adjustment method described above. Presuming that an effective adjustment procedure results in smooth year-to-year variations in ozone concentrations, this comparison shows that Davidson's temperature adjustment method appears to be more effective than the preliminary category-based adjustment method. Further analysis is needed to determine if a more effective category-based method can be developed.

CONCLUSIONS

On the basis of the results obtained in this study, it is evident that at least some of the day-to-day variations in meteorological conditions and in the spatial and temporal distribution of ozone in the SOCAB can be accounted for by the grouping of days into source history categories. Furthermore, we have shown that at least two of the source history categories originally identified by Zeldin (Eddy and Southern Route) are characterized by distinct and coherent spatial ozone patterns and meteorological conditions and that these differences characterize two different flow regimes:

Eddy days are characterized by southerly winds along the coast south of Palos Verdes, higher sea level pressure south of Los Angeles (at San Diego) and lower pressure to the east (at Las Vegas). These conditions are identical to those found by Mass and Albright (1989) to be characteristic of Catalina Eddy events. Basin wide ozone concentrations are lower under these conditions than on other days, apparently as a result of a deepening of the marine layer (thus resulting in increased cloudiness, increased dispersion and increased flow out of the basin of polluted air over mountain passes to the east) as suggested by Mass and Albright and confirmed by our data. Concentrations in the mountains and the San Fernando valley are higher relative to the basin average on these days than under other flow regimes.

Southern Route days are characterized by weak offshore pressure gradients with higher pressure to the north and east of Los Angeles and lower pressure to

the south (at San Diego). This pressure pattern results in a shallow marine layer and a delayed onset and a weakening of the sea breeze. Air mass residence time is thus increased over the more emission intensive western portions of the basin with minimal mixing, dispersion, and afternoon transport inland. As a result, some of the highest ozone concentrations of the season are observed under these conditions and concentrations in the heavily populated coastal and metropolitan regions are higher in both an absolute sense and relative to the basin-wide average than on other days.

A large proportion of days in each ozone season fail to exhibit some or all of the characteristics of either the Eddy or Southern Route patterns described above. Based on a preliminary analysis of the data used in our pilot study, Zeldin suggested that some of these days could be classified as belonging to a "Typical" source history category. However, we were unable to identify a consistent set of either spatial ozone patterns or meteorological conditions which could be used to characterize such a category. Zeldin also identified an "Offshore" category which appears to be primarily characterized by offshore pressure gradients, high temperatures, little or no sea breeze and relatively high ozone concentrations along the coast. Unfortunately, too few days of this type were found in the pilot study data set to allow for a definitive analysis of how such days might differ from those in the Southern Route category.

Although the Eddy and Southern Route patterns appear to identify distinctly different source history categories, there remains considerable variability in daily maximum ozone concentrations within each of these categories. Despite the fact that the means of the daily maximum ozone concentrations on days within each category are significantly different from one another in a statistical sense, only a small percentage of the variance in subregion average daily maximum ozone concentrations can actually be explained by the grouping of days into the Eddy, Southern Route and Uncertain categories. Temperature (especially the 850 mb temperature) appears to be the single most important factor in accounting for the remaining within-category variability.

Our attempts to use the knowledge of SOCAB source history categories gained in this study to adjust seasonal mean ozone trends to account for variations in meteorological conditions were not successful due to difficulties encountered in developing statistical models that do a good job of accounting for within-category variability.

RECOMMENDATIONS

Further analyses of the source history characteristics associated with ozone episodes are clearly needed before we can fully understand the complex relationships between meteorological conditions and ozone concentrations in the SOCAB. Therefore, we can at this time provide only preliminary guidance on the selection and application of meteorological trend adjustment procedures or on the development of episode selec-

tion criteria for photochemical modeling. With regard to trend adjustment methods, it is evident that the simple procedure developed by Davidson et al. (1985) is the best currently available method for making meteorological adjustments to trends in seasonal mean ozone where the mean is taken over all days (and thus all meteorological conditions). Nevertheless, differences between source history categories in unadjusted trends such as those shown in Figure 6-1 are of potential significance to air quality planners and should be examined further. In addition, further work on the development of more accurate regression models to be used for calculating adjusted within-category trends should be pursued.

With regard to episode selection criteria, the results of this study support those of previous studies in showing that high ozone concentrations can occur in the SOGAB under quite different meteorological scenarios. At this time we are only able to clearly identify two such scenarios (Eddy and Southern Route). Southern Route days exhibit characteristics favorable to the formation of high ozone concentrations, especially in the heavily populated central metropolitan areas. Ozone concentrations tend to be lower in most parts of the basin on Eddy days, although exceedances of the state and federal standards are still common, particularly at Newhall and Lake Gregory. Thus, the evaluation of Eddy and Southern Route days is important in the development of emission control strategies. In addition, a better understanding of the conditions associated with high ozone days that did not meet the Eddy or Southern Route meteorological criteria is needed. A case study approach in which the meteorological conditions on several such days are examined in detail would be an important first step.

Contents

Executive Summary	ii
List of Figures	xii
List of Tables	xvi
1 INTRODUCTION	1
2 DATA	3
Selection of Study Period	3
Selection of Aerometric Data	4
Selection of Ozone Summary Statistics	8
Selection of Meteorological Data	8
PART I: PILOT STUDY	
3 ANALYSIS OF SOURCE HISTORY CATEGORIES	14
Identification of Source History Categories by Expert Judgment	14
Temporal Distribution of Patterns	16
Meteorological and Aerometric Characteristics of Source History Categories	21
Development of Objective Ozone Pattern Criteria	23
Air Quality Characteristics of Ozone Patterns	29
Determination of Meteorological Signatures	34
Ozone-Meteorology Relationships Within Patterns	37
4 PILOT STUDY SUMMARY AND CONCLUSIONS	114
Summary	114
Conclusions	117

PART II: ADDITIONAL ANALYSES

5	FURTHER ANALYSES OF SOURCE HISTORY CATEGORIES	118
	Diurnal Profile Analysis	118
	Back Trajectory Analysis	125
6	DEVELOPMENT AND APPLICATION OF METEOROLOGICAL CATEGORIES	147
	Development of Meteorological Categories	147
	Relationship of Within-Category Ozone to Meteorology	151
	Ozone Trend Calculations	158
7	SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS	175
	Summary	175
	Conclusions	177
	Recommendations	178
	References	180

Figures

2-1	South Coast Air Basin monitoring stations operating during 1987 and associated subregions	6
3-1	Example data map for 5/3/83	43
3-2	Boxplots of relative subregion average daily maximum ozone concentrations for each source history category	44
3-3	Key for boxplots in Figures 3-2 and 3-4 through 3-22	53
3-4	Boxplots of relative subregion average daily maximum ozone concentrations	54
3-5	Boxplot of variable MAXTEMP for each source history category	63
3-6	Boxplot of variable WS0700 for each source history category	64
3-7	Boxplot of variable WS1000 for each source history category	65
3-8	Boxplot of variable SKY0700 for each source history category	66
3-9	Boxplot of variable CEIL0700 for each source history category	67
3-10	Boxplot of variable SKY1000 for each source history category	68
3-11	Boxplot of variable CEIL1000 for each source history category	69
3-12	Boxplot of variable LATHPRO7 for each source history category	70

3-13	Boxplot of variable LATHPR16 for each source history category	71
3-14	Boxplot of variable PGLAXWJF for each source history category	72
3-15	Boxplot of variable AWSLGB for each source history category	73
3-16	Boxplot of variable RWSLGB for each source history category	74
3-17	Boxplot of variable U0700 for each source history category	75
3-18	Boxplot of variable V0700 for each source history category	76
3-19	Boxplot of variable U1000 for each source history category	77
3-20	Boxplot of variable V1000 for each source history category	78
3-21	Boxplot of variable ULGB for each source history category	79
3-22	Boxplot of variable VLGB for each source history category	80
3-23	Boxplots of variable TMAXSBD for each source history category	81
3-24	Boxplots of variable T850 for each source history category	82
3-25	Boxplots of variable TMAXLAX for each source history category	83
3-26	Boxplots of variable DELTAT for each source history category	84
3-27	Boxplots of variable DELTAPG for each source history category	85

3-28	Boxplots of variable MAXDIFF1 for each source history category	86
3-29	Boxplots of variable MAXDIFF2 for each source history category	87
3-30	Key for boxplots in Figures 3-23 to 3-29	88
3-31	Distributions of inversion base height elevations from 1300 UTC LMU/UCLA sounding for days in each source history category	89
3-32	Classification tree illustrating relationship between ozone patterns and meteorological and aerometric variables	90
3-33	Boxplots of subregion average daily maximum ozone concentrations for each objectively determined ozone pattern	91
3-34	Percent of days in each ozone pattern on which subregion average daily maximum ozone concentration exceeds 20 pphm	100
3-35	Boxplots of pphm-hours per day above 9 pphm for each objectively determined ozone pattern	104
3-36	Classification tree illustrating relationship between objectively determined ozone patterns and meteorological variables	113
5-1	Diurnal variation of normalized hourly ozone concentrations	119
5-2	SOCAB wind observations, 11 July 1984, 1200 PST	126
5-3	SOCAB gridded wind field from station data interpolation, 11 July 1984, 1200 PST	127
5-4	Trajectories for Eddy days	129
5-5	Trajectories for Partial Eddy days	133
5-6	Trajectories for Typical days	136
5-7	Trajectories for Partial Southern Route days	140

5-8	Trajectories for Southern Route days	143
6-1	Seasonal mean ozone concentrations for each meteorological category: 3-year running averages	159
6-2	Three-year running average ozone trends with meteorological category based adjustment procedure applied (low ozone days excluded)	165
6-3	Three-year running average ozone trends with meteorological category based adjustment procedure applied (low ozone days included)	168
6-4	Seasonal mean ozone concentration trends with 850 mb temperature adjustment applied	170
6-5	Three-year running average seasonal mean ozone trend with 850 mb temperature adjustment applied	173

Tables

2-1	Period of record for stations assigned to each subregion in the South Coast Air Basin	5
2-2	Meteorological variables	9
2-3	Additional meteorological data	11
2-4	SCAQMD monitoring sites for which wind speed and direction data were obtained for the period 1984-1988.	13
3-1	Number of study days in each source history category by year	17
3-2	Contingency table showing number of source history categories by month, 1983-1985	18
3-3	Source history category sequence analysis	20
3-4	Distribution of days by source history category for each terminal node of the classification tree depicted in Figure 3-32	25
3-5	Objective criteria for identifying source history categories	27
3-6	Comparison of categories determined using the objective criteria in Table 3-5 with subjectively determined categories for 1983-1984 data	29
3-7	Comparison of categories determined using the objective criteria in Table 3-5 with subjectively determined categories for 1985 data	30
3-8	Number of days in each source history category on which subregion average daily maximum ozone concentration exceeds 20 pphm	33

3-9	Meteorological variables used in CART analysis	35
3-10	Distribution of source history categories by terminal node for the classification tree depicted in Figure 3-36	36
3-11	Meteorological variables used in regression analysis	38
3-12	Stepwise regression results	39
6-1	Meteorological criteria developed by Rosenbaum (1990) designed to identify days with Eddy and Southern Route type ozone patterns	149
6-2	Comparison of meteorological categories, ozone patterns, and Zeldin's classification scheme for Tuesdays, Wednesdays, and Thursdays, 1983-1985 and 1980-1988	150
6-3	Cross-validation results of CART regression tree analyses for subregion average daily maximum ozone 1980-1988	154
6-4	Relative importance of meteorological variables as determined by CART analyses	155
6-5	Number of times regression appear in top 10 best subset regressions	157
6-6	Percent of variance in subregion average daily maximum ozone concentrations that is accounted for by the partitioning of days into the Eddy, Southern Route, and Uncertain meteorological categories	162
6-7	Results of linear regression analyses for ozone trend adjustment procedure	164

I INTRODUCTION

Ozone concentrations in California's urban air basins exhibit considerable day-to-day and year-to-year variability. A large portion of this variability may be attributed to fluctuations in meteorological conditions. From the standpoint of air quality management, this intimate relationship between the weather and ozone has two principal consequences:

Meteorological conditions may vary significantly from one high ozone episode to the next. Because of the expense of obtaining data for applying photochemical grid models, it is only feasible to study the effects of emission control strategies on the basis of data collected for a small number of episodes. Since the effectiveness of a particular control strategy in reducing ozone concentrations will vary depending on weather conditions, it is difficult to predict the overall effect of a control strategy on the basis of results from just a few modeled episodes.

Fluctuations in the frequency and severity of weather conditions conducive to the formation of high ozone concentrations result in large differences in concentrations from one year to the next. These fluctuations effectively mask any underlying trend in ozone that may have resulted from changes in the geographical distribution and amount of precursor emissions.

The California Air Resources Board has long recognized the difficulties imposed on air quality management activities by the above consequences. To address these concerns, the ARB has contracted with Systems Applications to develop a series of methodologies designed to achieve the following goals:

Quantification of the effects of meteorological conditions on ozone concentrations in California air basins.

Classification of ozone episodes for control strategy modeling.

Calculation of meteorologically adjusted ozone trends.

These methodologies will allow air quality managers to deal with the problems noted above by providing them with the information needed to (1) make informed decisions on the representativeness of ozone episodes for purposes of emission control scenario

development, and (2) calculate year-to-year ozone trends that have been adjusted for meteorological fluctuations and thus more clearly indicate the effects of changes in precursor emission patterns.

During the planning stages of this study, it was determined that the best strategy for achieving the goals outlined above would be to develop a methodology for grouping days into categories such that days within the same category will have similar meteorological and air quality characteristics and thus be representative of one another in the sense described above. Given a method of identifying such "source history" categories, one could then select a limited number of days from each category for the purpose of modeling air quality responses to alternative emission control scenarios. In this way, model results would be available for a set of days that represent the range of episode conditions affecting the air basin under study. Furthermore, meteorologically adjusted air quality trends could be computed by calculating the change in ozone concentrations from one year to the next on days falling within the same source history category. Since such days are associated with similar meteorological conditions, any year-to-year changes in concentrations must be the result of changes in precursor emission levels or residual meteorological variability. The latter problem can be eliminated by developing relationships between meteorology and ozone for each category, and using these relationships to predict concentrations that would occur under standardized meteorological conditions.

In view of the uncertainties involved in attempting to develop a successful methodology of the type described, we first conducted a pilot study designed to determine the feasibility of the approach. Results of this pilot study, in which we examined data collected in the South Coast Air Basin (SOCAB) for the period 1983-1985, are presented in Part I of this report. Our analysis of these data indicated the potential presence of as many as four distinct groups of days in the SOCAB. Each group is associated with different combinations of ozone distribution patterns and meteorological conditions. This suggests that each group is associated with a coherent set of source history characteristics such that the mechanisms leading to the afternoon ozone peak in each portion of the basin (emission contributions from specific sources, mixing, transport, and chemical transformation) are the same for days within the same category.

However, evidence gathered during the pilot study in support of this finding is largely circumstantial. Therefore, additional analyses of ozone concentration patterns and meteorological conditions associated with each source history category were conducted as described in Part II of this report. These additional analyses, together with the results of a parallel study of SOCAB ozone patterns conducted by Rosenbaum (1990), provided sufficient information to allow for the classification of days into two distinct source history categories on the basis of values of several key meteorological indicator variables. These classifications were then applied to the calculation of meteorologically adjusted ozone concentration trends.

2 DATA

A detailed analysis of source history patterns requires an extensive collection of aerometric and meteorological data. For this reason, the well-studied South Coast Air Basin (SOCAB), with its large numbers of monitoring sites and extensive meteorological data sets, was selected for the pilot study. In the following sections we discuss the selection of time periods, monitoring sites, meteorological data, and daily and seasonal ozone summary statistics used in the pilot study.

SELECTION OF STUDY PERIOD

Although monitoring of ozone concentrations in the SOCAB has been conducted since the 1950s, the majority of high quality data has been collected since 1980. This year also corresponds to the widespread replacement of the older Potassium Iodide (KI) monitors with the more accurate and reliable UV photometric monitors. For this reason, we decided to confine our analysis to include only data collected since 1980.

To accomplish our goal of identifying groups of days with similar source history characteristics, it was necessary to identify a period of time over which precursor emissions could be assumed to have been fairly constant. This selection process consisted of several steps:

1. Selection by day-of-week. Because weekend and holiday emission patterns are substantially different from weekday patterns, and Monday and Friday traffic patterns differ from those on other weekdays, we eliminated weekends, Mondays, and Fridays from the database.
2. Selection by season. Because ozone levels tend to be lower during the winter months and vehicular emissions differ between the cold and warm weather months, only data for the months May-October were included in our analysis.
3. Selection of years. Since precursor emission rates vary over time as a result of automobile fleet turnover, implementation of new air quality regulations, population growth and other factors, it was necessary to limit our initial analysis to a period of three years or less. On the other hand, limiting the analysis to too short a time period (e.g., a single year) coupled with the restrictions imposed by items (1) and (2), would result in a very

small sample size. For this reason we decided to perform our pilot examination using two consecutive years of data with a third year held in reserve to test hypotheses. As described in the project Work Plan (Stoeckenius and Daly, 1989), we chose the years 1983-1985 for our analysis because this period does not seem to contain a preponderance of abnormal meteorology.

SELECTION OF AEROMETRIC DATA

As discussed in the Work Plan, ozone data are available at a large number of stations in the SOCAB. For our analysis we included only those stations using UV photometry which were not moved during the 1980-1988 study period and which were located in one of the nine subregions defined in Table 2-1 and Figure 2-1. Subregions were defined for the purpose of forming clusters of monitors experiencing similar ozone conditions. By performing our analysis on clusters of monitors rather than individual monitors, we were able to obtain more stable and precise ozone summary statistics. Monitoring subregions were identified using the following criteria:

Proximity of monitors to one another;

Similarity of locations with respect to terrain and land-use features; and

Knowledge gained from previous analyses of ozone climatology, transport and source-receptor relationships in the SOCAB.

The principal characteristics of each of the nine monitoring subregions we selected are described below:

1. North Coast--(West LA; Hawthorne, North Long Beach)
These sites characterize the coastal influences from about Palos Verdes northward. Sea breezes tend to carry pollutants east of these sites, leaving relatively low ozone levels behind.
2. South Coast--(Los Alamitos, Costa Mesa)
Similar to those in group 1, these sites are in a coastal environment, but located south of Palos Verdes. Also, these sites can be influenced at times by major power plant emissions, e.g., Los Alamitos by the Haynes/Alamitos power plants; Costa Mesa by the Huntington Beach power plant.
3. Metropolitan--(Los Angeles, Lynwood, Anaheim)
These sites are situated approximately 10-12 miles inland from the coast and are characterized by a predominance of mobile source emissions. Because of the presence of upwind sources between these sites and the coast, these sites typically have higher ozone levels than the coastal sites do, but considerably lower ozone levels than do most of the inland receptor sites.

TABLE 2-1. Period of record for stations assigned to each subregion in the South Coast Air Basin.

Subregion	Station	Site ID.	Years
1	North Long Beach	7000072	1980-1988
	W. Los Angeles (Robertson)	7000086	1980-1984
	W. Los Angeles (VA Hospital)	7000091	1984-1988
2	Los Alamitos	3000190	1980-1988
	Cosra Mesa	3000192	1980-1988
3	Anaheim	3000176	1980-1988
	Lynwood	7000084	1980-1988
	Los Angeles (North Main)	7000087	1980-1988
4	Burbank	7000069	1980-1988
	Reseda	7000074	1980-1988
5	Azusa	7000060	1980-1988
	Pasadena	7000088	1982-1988
	Glendora	7000591	1984-1988
6	La Habra	3000177	1980-1988
	Whittier	7000080	1980-1988
	Pico Rivera	7000085	1980-1988
7	Upland	3600175	1980, 1983-1988
	Fontana	3600197	1982-1988
	Pomona	7000075	1980-1988
8	Riverside	3300144	
	Redlands	3600192	1980-1986
	San Bernardino	3600194	1981-1986
9	Lake Gregory (Crestline)	3600181	1980-1988
	Newhall	7000089	1983-1988

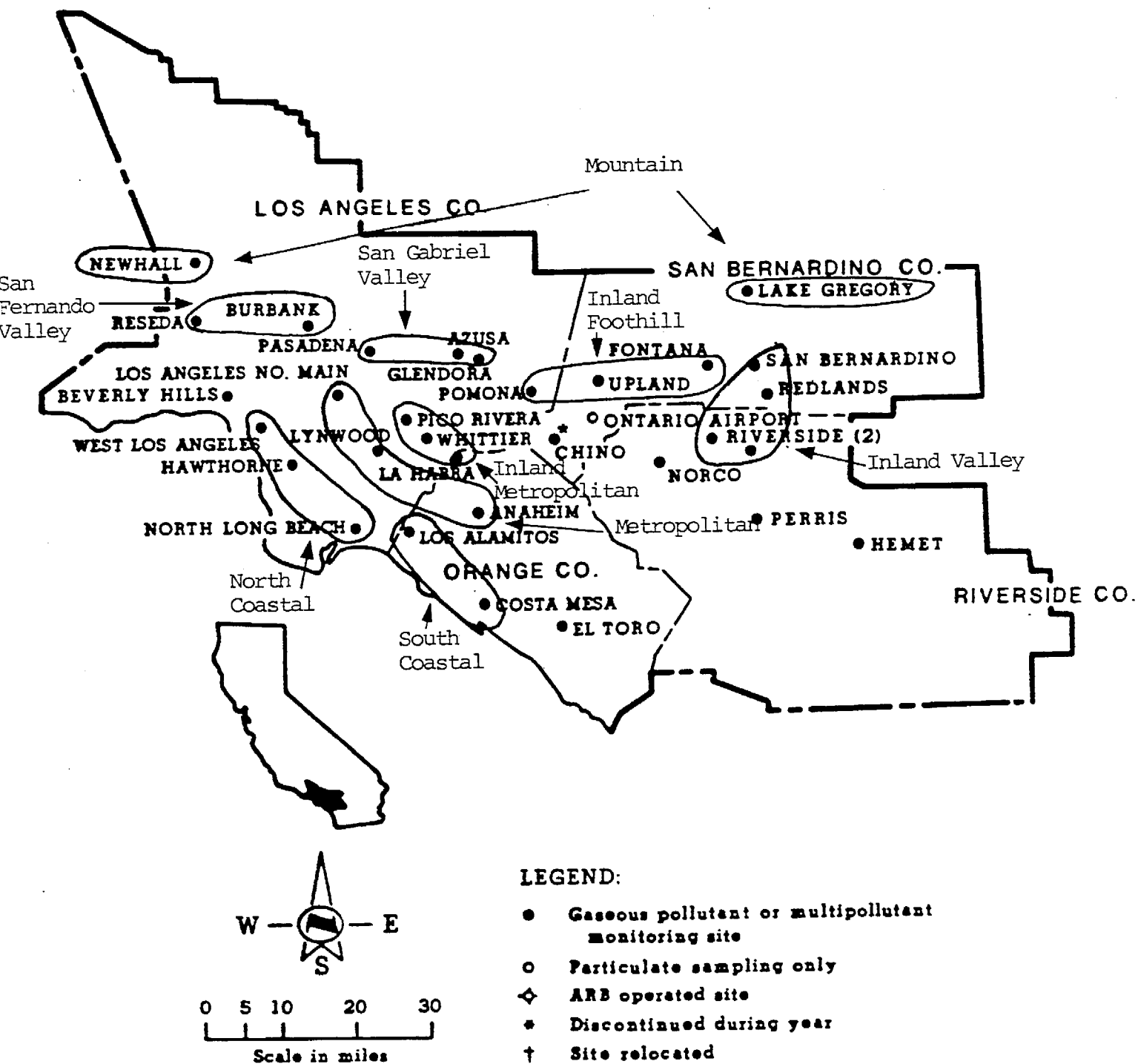


FIGURE 2-1. South Coast Air Basin monitoring stations operating during 1987 and associated subregions.

4. San Fernando Valley--(Burbank, Reseda)
The bifurcation of the sea breeze near the Hollywood Hills results in some pollution transport into the San Fernando Valley; however, in most cases, the bulk of the pollution is transported in the eastern branch of the flow. Hence, these sites typically have lower ozone levels than do other inland valley areas.
5. San Gabriel Valley--(Pasadena, Azusa, Glendora)
Currently noted as the basin's ozone "hotspot," the San Gabriel Valley typically records the highest daily ozone concentrations on most summer days. This situation most likely results from several factors: (1) the area is immediately downwind of the major Los Angeles emission source region; (2) temperatures are generally cooler than in the receptor areas to the east, indicating a greater marine influence with generally less vertical mixing; and (3) the ozone peaks typically occur between 1:00 and 3:00 p.m.--before emissions from the afternoon rush-hour traffic can cause "NO scavenging."
6. Inland Metropolitan--(Pico Rivera, Whittier, La Habra)
This grouping of monitors generally exhibits characteristics similar to those of the metropolitan subregion, except that ozone levels are slightly higher because of its more inland location. However, under weak or moderate offshore flow conditions, inland pollution transport can follow a trajectory farther south than normal, and sweep across this group of sites. In these instances, ozone levels can be similar to those observed in the San Gabriel Valley.
7. Inland Foothill--(Pomona, Upland, Fontana)
Farther inland than the San Gabriel Valley, these sites are major receptor sites for pollution transported eastward along the base of the San Gabriel Mountains. Because the ozone peaks typically occur between 2:00 and 4:00 p.m.--the time of maximum afternoon heating--greater vertical mixing usually helps to slightly reduce surface ozone levels compared to those in the San Gabriel Valley.
8. Inland Valley--(San Bernardino, Redlands, Riverside)
These sites represent the easternmost receptor sites in the basin. Occasionally, during weak Santa Ana wind conditions, these stations remain in relatively unpolluted air until late in the day, when sea-breeze-driven pollution finally sweeps across this subregion.
9. Mountain--(Newhall, Lake Gregory)
These elevated sites are located in northern portions of the basin. Although these sites are separated by approximately 80 miles, they exhibit similar air pollution climatologies.

SELECTION OF OZONE SUMMARY STATISTICS

To realize the statistical advantages of defining monitoring subregions, we based our analysis on the subregion averages of the daily maximum ozone concentrations at each monitor. In the pilot study, in which only data for the years 1983-1985 were examined, only valid daily maximums (i.e., those based on days with at least nine valid hourly measurements between 9:00 a.m. and 9:00 p.m. LST) were included in the subregion averages. Since only daily maximum and not hourly ozone values were obtained for the other study years (1980-1982 and 1986-1988), a different data completeness criterion (in which a daily maximum is not considered valid unless it is based on at least 22 valid hourly averages) was used for those analyses encompassing the entire 1980-1988 study period. No attempt was made to adjust the subregion averages to account for day-to-day changes in the number of monitoring stations reporting. In no case were we unable to calculate a subregion average due to a lack of valid daily maximum concentrations.

For the purposes of summarizing ozone concentrations in a subregion for a particular ozone season, we calculated the seasonal mean of the subregion average daily maximum concentrations along with the number of days on which the subregion averages exceeded 20 pphm, and the number of pphm-hours above 9 pphm. The latter (exposure) summary statistic was calculated by summing the difference between hourly ozone concentrations at each station in a subregion and 9 pphm, counting only those hours for which this difference is greater than zero. These individual-station pphm-hour values were then averaged over the stations in each subregion and summed over days.

SELECTION OF METEOROLOGICAL DATA

After reviewing the availability of various types of meteorological data in the SOCAB as described in the Work Plan, we initially focused on those variables we believed would provide the most information about both the source history characteristics (i.e., flow patterns) of each day as well as the degree to which daily meteorological conditions were conducive to ozone formation. A list of these variables appears in Table 2-2. Subsets of these variables were used at different stages of the pilot study as described in Section 3.

On the basis of the pilot study results, it was determined that additional meteorological data would be needed to complete our study. In particular, additional surface wind data were needed for air parcel trajectory calculations and more surface and upper air data were needed to better understand the meteorological signatures of each source history category. The additional data we obtained are summarized in Table 2-3 and described briefly next.

TABLE 2-2. Definitions of meteorological and ozone variables.

Name	Description
MAXTEMP	Daily maximum temperature at Thermal (deg. F).
TMAXSBD	Daily maximum temperature at San Bernardino (deg. F).
TMAXLAX	Daily maximum temperature at LAX (deg. F).
MAXDIFF1	TMAXSBD - TMAXLAX (deg. F).
MAXDIFF2	MAXTEMP - TMAXLAX (deg. F).
T850	850 mb temperature from 1300 UTC UCLA/LMU sounding (deg. C).
DELTAT	Temperature difference across inversion (Top - Base; deg. C).
DELTAPG	24-hour change in PGLAXWJF (mb).
WS0700	San Diego wind speed, 7:00 a.m. LST (kts).
WS1000	San Diego wind speed, 10:00 a.m. LST (kts).
SKY0700	Los Angeles airport (LAX) sky cover, 7:00 a.m. LST (tenths).
CEILO700	Ceiling height at LAX, 7:00 a.m. LST (feet, 99,999 = unlimited).
SKY1000	Los Angeles airport (LAX) sky cover, 10:00 a.m. LST (tenths).
CEIL1000	Ceiling height at LAX, 10:00 a.m. LST (feet, 99,999 = unlimited).
LATHPR07	Pressure gradient (LAX - Thermal), 7:00 a.m. LST (mb).
LATHPR16	Pressure gradient (LAX - Thermal), 4:00 p.m. LST (mb).
BASEHT	Inversion base height at UCLA/LMU, 1300 UTC (feet).
PGLAXWJF	Pressure gradient (LAX - Lancaster), 1300 UTC (mb).
AWSLGB	Daily scalar average wind speed, Long Beach (mph).
RWSLGB	Resultant wind speed, Long Beach (miles).

continued

TABLE 2-2. concluded

Name	Description
U0700	East-west component of unit wind direction vector, San Diego, 7:00 a.m. LST (negative value indicates wind from the west).
V0700	North-south component of unit wind direction vector, San Diego, 7:00 a.m. LST (negative value indicates wind from the south).
U1000	East-west component of unit wind direction vector, San Diego, 10:00 a.m. LST (negative value indicates wind from the west).
V1000	North-south component of unit wind direction vector, San Diego, 10:00 a.m. LST (negative value indicates wind from the south).
ULGB	East-west component of resultant daily unit wind direction vector, Long Beach (negative value indicates wind from the west).
VLGB	North-south component of resultant daily unit wind direction vector, Long Beach (negative value indicates wind from the south).
RATNET<n>	Ratio of the subregion average daily maximum ozone concentration to the basin-wide average daily maximum ("relative subregion average daily maximum concentration") for the nth subregion.

TABLE 2-3. Additional meteorological data.

Parameter		Stations
<u>Upper Air</u>		
Temperature	} 850, 500 mb 0400, 1600 PST	Oakland, CA
Wind direction		Ely, NV
Wind speed		Vandenburg, CA
Geopotential height		Desert Rock, NV San Diego, CA
<u>Surface</u>		
Ceiling height	} 0700, 1000, 1300, 1600 PST	LAX
Dew point		San Diego
Opaque sky cover		Las Vegas
Relative humidity		
Wind direction		
Wind speed		
Daily maximum temperature		
Sea level pressure		
Resultant wind speed	} Daily averages	LAX
Resultant wind direction		Long Beach
Daily maximum temperature		Burbank
Average relative humidity		Ontario
		El Toro
	Downtown Los Angeles	
	Riverside	
	San Bernardino	
	Lancaster	

Surface Data

For purposes of the trajectory calculations, additional surface wind data were obtained from the South Coast Air Quality Management District (SCAQMD). SCAQMD collects and archives 5-minute average wind speed and direction observations at each of the air quality monitoring sites it operates. Due to a data processing problem at SCAQMD, hourly resultant wind directions to the nearest degree are not available prior to October 1987. However, wind speed and direction observations representing the last minute of each hour are available prior to this time. In addition, the data processing problem did not affect the 16-point scale resultant wind direction calculation. We therefore obtained the 16-point hourly prevailing wind data for use in the back trajectory analysis for all hours and days of the week for the period 1984-1988. A list of the stations for which wind observations were obtained appears in Table 2-4.

We also obtained from SCAQMD daily values of resultant wind speed and direction, maximum temperature and average relative humidity for several SOCAB sites as indicated in Table 2-3.

In addition to the wind data just described, surface observations collected once every three hours at LAX, San Diego and Las Vegas were obtained from the National Climatic Data Center for the period 1980-1988. Of particular importance in these data are the wind speed and direction, temperature and sea level pressure observations.

Upper Air Data

Wind, temperature and geopotential heights at 850 and 500 mb were obtained from twice daily soundings conducted at the four observing stations closest to Los Angeles: Vandenburg Air Force Base, San Diego Montgomery Field, Oakland International Airport and Mercury Desert Rock (near Las Vegas). These data were obtained from the ATAD North America upper-air data tapes produced by the National Climatic Data Center.

TABLE 2-4. SCAQMD monitoring sites for which wind speed and direction data were obtained for the period 1984-1988 (not all sites available in all years).

<u>Los Angeles County</u>	<u>Riverside County</u>
Azusa	Palm Springs
Burbank	Hemet
Long Beach	Riverside
Reseda	Perris
Pomona	Banning ALLES
Lennox	Norco
Whittier	
Lancaster	<u>San Bernardino County</u>
Lynwood	
Pico Rivera	Upland
West Los Angeles	Crestline
Los Angeles	Redlands
Pasadena	San Bernardino
Santa Clarita	Fontana
West Los Angeles VA	Chino
Glendora	
<u>Orange County</u>	
Anaheim	
La Habra	
El Toro	
Los Alamitos	
Costa Mesa	

PART I: PILOT STUDY

3 ANALYSIS OF SOURCE HISTORY CATEGORIES

As noted in the Introduction, our general approach is to develop a method for grouping days into categories such that days within the same category will have similar meteorological and air quality characteristics and thus be representative of one another with respect to the relationships between ozone receptor sites and source areas and ozone formation mechanisms. To develop an acceptable categorization scheme, we adopted a two-phase approach:

Phase I: Examine both meteorological and aerometric data over a period of time during which precursor emission levels can be assumed to have been nearly constant. Look for distinct, recurring patterns of pollutant concentrations and meteorological conditions. Group days into source history categories according to these patterns.

Phase II: Identify the unique meteorological aspects (i.e., the meteorological "signatures") of each source history category identified in Phase I. This will allow for the classification of days into these categories even for time periods in which precursor emissions differ significantly from those during the period examined in Phase I.

It was necessary to use both meteorological and aerometric data in the first phase because the available meteorological data are insufficient to define the complex time-varying three-dimensional atmospheric structure responsible for ozone formation and transport, making it impossible to identify the characteristics of each source history category. Once the source history categories were identified in Phase I, we searched for the unique meteorological profiles of each category as called for in Phase II.

We began our analysis by summarizing the available meteorological and ozone data for each day in the study period and identifying the most likely principal source history categories based on expert knowledge of the ozone climatology of the SOCAB. This was followed by an examination of the principal meteorological and aerometric characteristics of each category.

IDENTIFICATION OF SOURCE HISTORY CATEGORIES BY EXPERT JUDGMENT

Our analysis of ozone episode patterns in the SOCAB is based on the premise that such patterns are best identified by examining both meteorological conditions and

ozone concentration patterns. We therefore prepared data maps of the SOCAB for each day in the study period for the years 1983-1985 that show the subregion average daily maximum ozone concentrations and the values of several key meteorological variables. An example of such a map appears in Figure 3-1. (All figures appear at the end of this section.) The complete set of maps is available upon request. Analysis of these maps by Mel Zeldin of the South Coast Air Quality Management District resulted in the identification of eight source history categories defined as follows:

Typical Pattern (Category #3): Days in this pattern are characterized by westerly or northwesterly coastal winds that transport ozone and precursors through the San Gabriel Valley and into the eastern portions of the basin.

Eddy Pattern (Category #1): Days in this pattern are characterized by coastal winds with a pronounced southerly component that produce a greater transport of ozone and precursors into the San Fernando Valley and upslope into the Mountain subregion than occurs on Typical Pattern days.

Southern Route Pattern (Category #6): This pattern includes days with a more northerly wind flow aloft that tends to reduce the northeasterly flow over the San Gabriel Valley subregion and push ozone and precursors further into Orange County.

Offshore Pattern (Category #7): This pattern includes days in which a weak offshore (i.e., negative) coastal to inland pressure gradient reduces the eastward extent and vigor of the seabreeze circulation, resulting in ozone levels that are higher in the coastal and metro subregions and lower in the easternmost portions of the basin than on Typical Pattern days.

Three additional categories representing combinations of the above patterns were also identified:

Partial Eddy Pattern (Category #2): Days in this category exhibit spatial ozone distributions that are similar to those observed in the Eddy Pattern but with slightly lower values in the San Fernando Valley and Mountain subregions. However, the southerly component winds characteristic of the Eddy Pattern are not present on these days.

Typical Pattern with Eddy Winds (Category #4): A few days were identified in which coastal winds had a definite southerly component but the distribution of relative ozone concentrations was similar to those found on Typical Pattern days.

Partial Southern Route (Category #5): Days in this category exhibit some of the meteorological characteristics of Southern Route days (category #6), with a weak push of pollutants into northern Orange County but with sufficient

afternoon westerly winds to produce the usual basin-wide maximum concentrations in the San Gabriel Valley and Inland Foothill and Inland Valley subregions characteristic of Typical Pattern days. This scenario results in higher relative concentrations in the Inland Metro subregion than are observed on Typical Pattern days.

Finally, all days in which no subregion average daily maximum ozone concentration in excess of 8 pphm occurred were assigned to an eighth, low ozone, category.

The subjective expert judgment analysis described above was intended to result in the identification of source history categories that represent unique combinations of meteorological scenarios and ozone concentration patterns. Although the general characteristics of these categories are known and were used as the basis for the initial classification, detailed quantitative descriptions of the meteorological and air quality aspects of each pattern are needed to confirm the unique nature of each category and to provide additional information on what makes days in one category different from those in another. This information is also needed for the development of a more rigorous (i.e., an objective) procedure for classifying days into source history categories.

We first analyzed the distribution of categories over time to determine their overall prevalence and distribution within the ozone season. This was followed by a comparison of the distributions of meteorological parameters and relative ozone concentrations within each category. On the basis of this information, we developed a series of criteria that can be used to objectively assign any given day to one of the source history categories on the basis of meteorological data and relative ozone concentrations. With the days reclassified according to these objective criteria, we then compared distributions of the absolute magnitudes of ozone concentrations from one source history category to the next.

TEMPORAL DISTRIBUTION OF PATTERNS

As a first step towards understanding the characteristics of the source history categories described in the preceding section, we examined their temporal distribution. Table 3-1 indicates the number of days assigned by Mel Zeldin's subjective analysis to each of the categories just described. The Eddy Pattern is the most prevalent, containing 50 percent more days than the next most numerous category (#2). Only 38 days over the three-year period exhibited both the aerometric and meteorological characteristics of a "typical" day, although numerous other days (i.e., those in categories #2, #4 and #5) exhibited some of these characteristics. The Offshore Pattern (category #7) appeared on only four days in one of the three years, and categories #4 and #6 were also quite uncommon.

Some of the ozone patterns described above appear more frequently during certain portions of the May - October season we examined. Table 3-2 displays the number of

TABLE 3-1. Number of study days in each source history category by year.

Category	Year			Total 1983-1985
	1983	1984	1985	
1	26	26	16	68
2	12	17	12	41
3	10	10	18	38
4	1	2	9	12
5	11	7	10	28
6	5	5	1	11
7	4	0	0	4
8	9	13	14	36

TABLE 3-2. Contingency table showing number of source history categories by month, 1983-1985.

Category	Month					
	May	June	July	August	September	October
1	18	19	10	9	5	7
2	7	10	12	3	6	3
3	8	1	6	17	4	2
4	2	0	4	3	0	3
5	1	4	6	6	7	4
6	3	0	1	0	4	3
7	0	0	0	2	0	2
8	3	4	0	1	11	17

days by month for the years 1983 - 1985 for each category. If the distributions of categories across months were random, we would expect to see a roughly equal number of days in each month in each row of the table. The values in Table 3-2 indicate that this is not the case as confirmed by a Chi-Square test (Chi-Square = 115.2 with 35 degrees of freedom which is significant at the 99.9 percent confidence level).^{*} Roughly defining May - June as the early season, July - August as mid-season, and September - October as the late season, Table 3-2 indicates that the Eddy Pattern (category #1) is more prevalent during the early season, while Typical Pattern days occur more frequently during the mid-season when the onshore winds tend to be the strongest. Southern Route days appear to be more prevalent during the early and late season, when the driving mechanism for the sea-breeze regime is weaker. As expected, low ozone days (category #8) are also most frequent during the early and late season months. Thus, the distribution of categories across the May-October ozone season is consistent with their meteorological characteristics.

Since some of the ozone patterns are closely related to one another (e.g., categories #1, #2 and #3) while other are not (e.g., categories #3 and #7), we hypothesized that patterns may characteristically occur in certain sequences. We therefore searched for pattern sequences by counting the number of days on which one pattern follows another. Since our study period only included Tuesdays, Wednesdays and Thursdays, we counted sequences from Tuesday to Wednesday and Wednesday to Thursday. The resulting contingency table is presented as Table 3-3. Numbers in parentheses indicate the expected counts in each cell of the table under the hypothesis that the category assigned to a particular day is independent of the category assigned to the previous day. Although a Chi-square statistic significant at the 99.5 percent confidence level was obtained for this table, the statistic is not reliable since so few data are available for some of the categories. Nevertheless, examination of the table reveals that higher than expected counts were obtained for all of the entries representing a continuation of the same ozone pattern from one day to the next (i.e., all cells on the main diagonal), with the sole exception of category #7 for which very little data are available. Beyond this tendency towards persistence, no other preferred pattern sequences are readily apparent; however, given the limited amount of data available, they cannot be ruled out altogether.

^{*} Strictly speaking, this Chi-Square test may not be valid since over half of the cells in Table 3-2 have expected counts of less than 5. However, both the large value of the statistic and the large differences in cell frequencies between months for the more numerous categories suggest that the distribution is not random.

TABLE 3-3. Source history category sequence analysis (1983-1984 data; numbers in parentheses indicate expected cell counts assuming independence).

Category of Tuesday or Wednesday	Category of Following Wednesday or Thursday							
	1	2	3	4	5	6	7	8
1	14 (10.3)	7 (5.7)	3 (4.0)	*	3 (3.1)	0 (1.4)	1 (0.9)	2 (4.6)
2	7 (7.2)	0 (4.0)	2 (2.8)	*	2 (2.2)	0 (1.0)	0 (0.6)	3 (3.2)
3	3 (5.1)	4 (2.9)	6 (2.0)	*	2 (1.6)	0 (0.7)	0 (0.4)	0 (2.3)
4	1 (1.0)	0 (0.6)	1 (0.4)	*	0 (0.3)	0 (0.1)	0 (0.1)	1 (0.5)
5	5 (3.4)	0 (1.9)	1 (1.3)	*	2 (1.0)	2 (0.5)	0 (0.3)	0 (1.5)
6	4 (3.4)	2 (1.9)	0 (1.3)	*	1 (1.0)	3 (0.5)	0 (0.3)	0 (1.5)
7	2 (1.0)	0 (0.6)	0 (0.4)	*	0 (0.3)	0 (0.1)	0 (0.1)	1 (0.5)
8	0 (4.5)	0 (2.5)	1 (1.7)	*	1 (1.4)	0 (0.6)	2 (0.4)	9 (2.0)

* All three days in category 4 were Tuesdays.

METEOROLOGICAL AND AEROMETRIC CHARACTERISTICS OF SOURCE HISTORY CATEGORIES

We examined the meteorological and aerometric characteristics of each source history category by calculating percentiles of the frequency distributions of the relative ozone concentrations and various meteorological variables within each of them. This analysis was performed only on the 1983-1984 data set. Data from 1985 was held in reserve to serve as a test sample with which to check our hypotheses concerning the principal characteristics of each category.

Relative subregion average daily maximum ozone concentrations were calculated for each day by dividing the subregion average daily maximum concentration by the basin-wide average daily maximum concentration. This normalization puts days with different absolute (basin-wide) ozone formation, such as may occur as a result of seasonal changes or temperature differences, on the same footing. Boxplots showing the mean and various percentiles of the relative ozone concentrations by category for each subregion are presented in Figure 3-2.¹ A key to these plots is presented in Figure 3-3. To facilitate comparisons across subregions, the same boxplots are presented for all subregions by category in Figure 3-4.

We also prepared boxplots showing percentiles of the distributions of several key meteorological variables within each category (Figures 3-5 to 3-29). A key to the plots in Figures 3-5 to 3-22 appears in Figure 3-3, and a key to the plots in Figures 3-23 to 3-29 appears in Figure 3-30. The variables we examined are defined in Table 2-2.

In addition, we compared the distributions of inversion base height (BASEHT) at UCLA/LMU (1300 UTC) within each source history category. The base height was broken down into four categories for this purpose as shown in Figure 3-31.

Based on the boxplots in Figures 3-2, 3-4, 3-5 to 3-29, and the bar chart in Figure 3-31, certain key characteristics of each source history category can be identified:

Typical Pattern (Category #3): Highest relative ozone readings on days in this category occur in the San Gabriel Valley, Inland Foothill, and Inland Valley subregions, with the lowest levels in the North and South Coast subregions. This ozone pattern is consistent with the predominantly westerly or north-westerly coastal winds, onshore pressure gradients, and cool coastal and warm inland temperatures found on these days.

¹ Relative ozone concentrations for each subregion are referred to as RATNET<n> where <n> is the subregion number as indicated in Table 2-1.

Eddy Pattern (Category #1): The distinguishing meteorological feature of eddy days is a southerly or southwesterly wind at San Diego and Long Beach. These winds push pollutants north into the San Fernando Valley and up the adjacent mountain slopes, producing higher relative ozone readings in the mountain and San Fernando Valley than on Typical Pattern days. In addition, slightly lower relative concentrations occur in the San Gabriel Valley.

Southern Route Pattern (Category #6): This category is characterized by near zero LAX - Thermal and negative LAX - Lancaster pressure gradients, mostly clear mornings at LAX (SKY0700 near zero), and surface-based morning inversions (BASEHT \leq 100 ft) at LAX/LMU. San Diego winds are primarily out of the northwest on these days, suggesting the existence of a generally north or northwest flow over the region. This meteorological scenario produces a pattern of relative ozone concentrations that is markedly different from that observed on Typical Pattern days. Higher relative concentrations occur in both the North and South Coast subregions as well as the Metro and Inland Metro subregions while the Inland Foothill, Inland Valley, and Mountain subregions experience lower relative concentrations on these days.

Offshore Pattern (Category #7): The four days during 1983 - 1984 assigned to this pattern are characterized by relative ozone concentrations in the North and South Coast and San Fernando Valley subregions that are higher than on Typical Pattern days, whereas Inland Foothill, Inland Valley and Mountain areas had lower relative concentrations. These conditions are the result of weak, negative (offshore) coast-to-desert pressure gradients, together with cooler desert temperatures (as measured at Thermal) that greatly reduce the strength and inland extent of the afternoon sea breeze.

Partial Eddy Pattern (Category #2): Meteorological conditions on days in this category are similar to those observed on Typical Pattern days, with the exception that winds at Long Beach tend to have a slightly larger southerly component. Relative ozone concentrations for this pattern at San Fernando Valley, Inland Metro, North Coast, and Inland Foothill are intermediate between those observed on Eddy and Typical Pattern days. Other subregions, including the Mountain subregion, experience relative concentrations similar to those observed on Eddy Pattern days.

Typical Pattern with Eddy Winds (Category #4): The three days in 1983 - 1984 assigned to this category exhibit coastal winds with a southerly component but with relative ozone concentrations more characteristic of Typical Pattern days. Early morning (7:00 am) winds at San Diego on these days are light and variable, suggesting that transport winds are not well defined.

Partial Southern Route Pattern (Category #5): Days in this pattern exhibit a greater frequency of clear mornings at LAX and shallower inversion bases than do Typical Pattern days. In addition, average and median values of TMAXSBD

and T850 are higher for days in this category than for any other category. In other respects, meteorological conditions on these days are similar to those on Typical Pattern days. Relative ozone concentrations at the Mountain and Inland Metro subregions are intermediate between those observed on Typical and Southern Route pattern days, while concentrations at other subregions are similar to those in the Typical Pattern.

Low Ozone Pattern: Key features of this pattern are the low 850mb and San Bernardino temperatures (T850 and TMAXSBD) and the uniformity of ozone concentrations across the basin.

The results just described indicate that there are significant differences in meteorological characteristics and spatial ozone distributions between the source history categories. This is particularly true of the four primary categories (Eddy, Typical, Southern Route, and Offshore), each of which is associated with quite different sets of conditions. As expected, the Partial Eddy Pattern, Typical Pattern with Eddy Winds, and Partial Southern Route Pattern each represent combinations of the characteristics of two of the primary categories although not enough days were classified as Typical Pattern with Eddy Winds to allow for any generalizations concerning how this pattern fits in with the others. The Partial Eddy Pattern appears to represent the middle ground between Eddy and Typical days so that there is no clearcut gap between these two patterns but rather a continuous transition from Eddy to Partial Eddy to Typical conditions. A particular day may fall anywhere within this spectrum; no "preferred" modes are evident from our analysis. Much the same can be said of the Partial Southern Route Pattern, which shares characteristics of both the Typical and Southern Route patterns.

DEVELOPMENT OF OBJECTIVE OZONE PATTERN CRITERIA

To gain a better understanding of the principal features of, and relationships between, the source history categories described above and to provide a way of quickly and unambiguously categorizing daily data, we attempted to develop a series of objective criteria that can be used to identify the source history category in which a particular day would most likely have been classified by a subjective analysis. These criteria are based on both the analyses of relative ozone concentrations and meteorological variables discussed above.

CART Analysis of Subjectively Determined Categories

As an initial step towards the development of objective criteria, we grew a classification tree using the CART methodology (see Appendix A of the Workplan for a description of CART) in which the subjectively determined category for a particular day served as the dependent variable, and the relative ozone and meteorological variables served as the predictor variables. The resulting classification tree is

depicted in Figure 3-32. The first split is on the 10:00 a.m. wind direction at San Diego (WD1000). All days with wind directions greater than 245° are moved to the upper node, while the remaining days are moved to the lower node. Since there are almost no occurrences of east component winds, the set of all days with WD1000 greater than 245° includes only days with north or northwest winds. The remaining splits in the tree are based on the values of the relative subregion average daily maximum ozone concentrations.

Table 3-4 summarizes the distribution of source history categories for each of the terminal nodes of the classification tree depicted in Figure 3-32. Examination of the columns in this table reveals that the classification tree does a reasonably good job of isolating the conditions associated with each category. For example, of the 18 days assigned to category #5, 16 were placed in the 4th terminal node. This node contains only two other days, one from category #3 and one from category #8. Thus, the conditions associated with the 4th terminal node ($WD1000 > 245$, $RATNET9 \leq 1.24$, $RATNET6 > 1.0$, $RATNET1 \leq 0.7$ or, in other words, days with northerly or northwesterly winds, relatively low ozone concentrations in the Mountain and North Coast subregions and relatively high concentrations in the Inland Metro subregion) can be used to identify days that are assigned to category #5 by the subjective procedure with good accuracy.

In CART, each terminal node is assigned to the category representing the greatest fraction of days in the node. In this context, misclassifications represent instances in which the category of a particular day falling in the node differs from the category assigned to the node. For the 1983 - 1984 data used to grow the tree in Figure 3-32, the misclassification rate is 22 percent. In other words, 22 percent of days are incorrectly classified by the tree. Since the tree itself was constructed on the basis of the 1983 - 1984 data, it is possible that this "resubstitution" misclassification rate is artificially low and that the true rate (which would be obtained by using the tree to classify a large number of independently selected days) is higher. CART provides a cross-validation estimate of the true misclassification rate (see Breimen et al., 1984, for details). For the tree in Figure 3-32, this estimate is 32 percent.

Looking again at Table 3-4, it appears that a number of the misclassifications discussed above result from the inclusion of low ozone days (category #8) in terminal nodes other than node 2 (which is made up of mostly low ozone days). This is not surprising given that the meteorological and relative ozone variables included in the CART analysis were chosen for their ability to categorize days on the basis of source history categories and not on the basis of absolute ozone concentration levels. Terminal node 2 consists of days with the southerly component winds characteristic of eddy days but without the characteristically low South Coast subregion relative ozone concentrations. Most of these days are low ozone days, but some days from other categories are also included. Other low ozone days (which occur under a variety of wind directions) are spread throughout the remaining nodes of the tree. Because of their relatively small numbers within each node, it was not possible for CART to determine statistically significant splits to further isolate these days. This

TABLE 3-4. Distribution of days by source history category for each terminal node of the classification tree depicted in Figure 3-32.

Terminal Node	Category								Total Days
	1	2	3	4	5	6	7	8	
1	41	0	0	1	0	0	1	1	44
2	6	1		2	0	0	1	16	26
3	1	2	15	0	1	0	2	2	23
4	0	0	1	0	16	0	0	1	18
5	1	0	0	0	1	10	0	0	12
6	3	26	4	0	0	0	0	2	35

is also true for some other categories, such as category #4, for which too few days were available on which to base a good set of splitting criteria.

Another source of misclassifications in our tree stems from a lack of precise demarcation between categories #1 and #2 and #2 and #3. Nearly a third of the category #3 days are misclassified as category #2 (terminal node 6).

From the above discussion, it appears that a primary reason the classification tree we developed results in misclassifications is that the data are split on the basis of the value of only one variable at a time. Thus, combinations of conditions can only be identified through a series of splits that often leave too little data to be reliably analyzed. A better classification system could be devised by simultaneously considering combinations of variables rather than examining them one at a time. Nevertheless, the above CART analysis can be used to guide us towards reasonable objective definitions of each source history category.

Identification of Objective Classification Criteria

In developing a set of objective classification criteria, we attempted to capture the principal features of each category without including a large amount of extraneous detail that may turn out to apply only to the particular data set analyzed. For simplicity, we only considered simultaneous criteria for each category; a decision tree consisting of hierarchical sets of criteria was not considered. The criteria for each pattern that resulted are presented in Table 3-5. Eddy days (category #1) are identified by winds with a southerly component at San Diego and Long Beach and high relative ozone concentrations in the Mountain subregion. Partial Eddy days (category #2) are identified by a switch to northwesterly winds at San Diego but continued high ozone in the mountains. Typical days (category #3) are identified by northwesterly winds, low Mountain and Inland Metro ozone and onshore pressure gradients. Typical Pattern with Eddy wind (category #4) days are identified by more southerly winds, but continued lower Mountain ozone and onshore pressure gradients. Partial Southern Route days (category #5) have northwesterly winds and relatively high Inland Metro ozone readings with low ozone in the mountains and along the northern coast. Southern Route days (category #6) also have northwesterly winds but with an offshore pressure gradient between LAX and the northern (high) deserts and an onshore gradient between LAX and the eastern (low) deserts. They are also differentiated from Partial Southern Route days by higher ozone readings in the North Coast subregion. Offshore days (category #7) are identified by offshore pressure gradients between LAX and both Thermal and Lancaster. Low Ozone days (category #8) are defined as any day in which no subregion average daily maximum ozone concentration exceeds 80 ppb.

TABLE 3-5. Objective criteria for identifying source history categories.

Category	Criteria*
1 Eddy	$100^{\circ} \leq \text{WD1000} < 245^{\circ}$ and $150^{\circ} \leq \text{RWDLGB} \leq 240^{\circ}$ and $\text{RATNET9} > 1.25$
2 Partial Eddy	$245^{\circ} \leq \text{WD1000} \leq 360^{\circ}$ and $\text{RATNET9} \geq 1.25$
3 Typical	$245^{\circ} \leq \text{WD1000} \leq 360^{\circ}$ and $\text{RATNET9} < 1.25$ and $\text{RATNET6} < 1.02$ and $\text{LATHPR16} \geq 0$ and $\text{PGLAXWJF} \geq 0$
4 Typical Pattern-eddy winds	$100^{\circ} \leq \text{WD1000} < 245^{\circ}$ and $\text{RATNET9} < 1.25$ and $\text{LATHPR16} \geq 0$ and $\text{PGLAXWJF} \geq 0$
5 Partial Southern Route	$245^{\circ} \leq \text{WD1000} \leq 360^{\circ}$ and $\text{RATNET9} < 1.25$ and $\text{RATNET6} \geq 1.02$ and $\text{RATNET1} \leq 0.7$
6 Southern Route	$245^{\circ} \leq \text{WD1000} \leq 360^{\circ}$ and $\text{LATHPR16} > 0$ and $\text{PGLAXWJF} < 0$ and $\text{RATNET1} > 0.7$
7 Offshore	$\text{LATHPR16} < 0$ and $\text{PGLAXWJF} < 0$
8 Low Ozone	All subregion average daily maximum ozone ≤ 8 pphm

* Variable names are defined in Table 2-2.

Comparison of Objectively and Subjectively Determined Categories

Table 3-6 compares categories assigned to days in 1983 - 1984 by the objective criteria in Table 3-5 to the subjectively determined categories. Of the 157 days during this period, 21 did not meet the objective criteria for any pattern. For most of the categories, the objective and subjective methods resulted in the same classifications. A notable exception is the subjectively determined Typical Pattern days, many of which did not meet the objective criteria for any pattern, and, of the remainder, almost half were assigned to category #2 by the objective procedure. Since categories #2 and #3 are quite similar in many respects, the mixup between these two categories is not surprising. The relatively high proportion of days not meeting the objective criteria for the Typical Pattern is probably a reflection of the large number of conditions such days must meet and suggests that the subjective criteria may be overly specific for this category.

To test the ability of the objective method to assign days to categories identical to those assigned by the subjective method on an independent data set, we used the 1985 data as a test sample. Table 3-7 compares the subjective and objective classifications. Categories #1 and #2 were similarly identified by the two methods. Of the 18 subjectively determined category #3 days, 10 were identified as belonging to this category by the objective procedure, with the remainder split between categories #2, #5 and #6. The two approaches differed most notably on the classification of category #5 and #6 days. Despite these differences, the objective classification procedure should result in the identification of groups of days with source histories that are very similar to those identified by the subjective analysis.

Having developed a suitable objective procedure for identifying source history categories that relies in part on the relative magnitudes of ozone concentrations in each subregion, we now examine differences between source history categories in the absolute magnitudes of ozone concentrations in each subregion.

AIR QUALITY CHARACTERISTICS OF OZONE PATTERNS

Up to this point, we have dealt only with relative ozone concentrations in describing the characteristics of source history categories. Since each category is associated with a different range of meteorological conditions, we can expect significant differences to exist in the absolute magnitudes of ozone concentrations on days falling into different categories. In particular, some categories may regularly be associated with the highest concentrations that occur in the SOCAB while others may represent days with much lower ozone levels. Those categories associated with extreme ozone episodes will be of particular importance in the development of emission control scenarios.

TABLE 3-6. Comparison of categories determined using the objective criteria in Table 3-5 with subjectively determined categories for 1983-1984 data.

Objective Category	Subjective Category							
	1	2	3	4	5	6	7	8
Missing	10	1	8	1	0	0	1	0
1	35	0	0	1	0	0	0	0
2	3	26	4	0	0	0	0	0
3	0	0	6	0	0	0	0	0
4	1	0	0	1	0	0	0	0
5	0	0	2	0	17	0	0	0
6	1	2	0	0	1	8	0	0
7	1	0	0	0	0	2	3	0
8	1	0	0	0	0	0	0	22

TABLE 3-7. Comparison of categories determined using the objective criteria in Table 3-5 with subjectively determined categories for 1985 data.

Objective Category	Subjective Category							
	1	2	3	4	5	6	7	8
Missing	3	0	1	3	2	0	0	1
1	11	0	0	1	0	0	0	0
2	1	12	2	0	0	0	0	0
3	0	0	10	2	0	0	0	0
4	1	0	0	3	0	0	0	0
5	0	0	3	0	4	0	0	0
6	0	0	2	0	3	1	0	0
7	0	0	0	0	1	0	0	0
8	0	0	0	0	0	0	0	13

to the next. With the exception of the San Fernando Valley subregion, this analysis showed that, with better than 99.9 percent certainty, the averages are not all equal (only categories #1-7 were included in this analysis). For the San Fernando Valley subregion, the null hypothesis could only be rejected at the 90 percent confidence level. The analysis further showed that the source history categories account for 39 to 52 percent of the variance in the North and South Coastal, Metropolitan and Inland Metropolitan subregions, and 22-30 percent in the subregions lying further inland. Only 6 percent of the variance in the San Fernando Valley is accounted for.

Exceedance Days

To identify the extent to which each category is associated with peak ozone events, we calculated the percentage of days in each category on which the subregion average daily maximum ozone concentrations exceed 20 pphm. Results for each subregion are presented in Table 3-8 and Figure 3-34 and are generally consistent with those for the daily maximum concentration presented above. Exceedances in the North and South Coast subregions occur very infrequently and only under the Southern Route Pattern. Some exceedances also occur on Partial Southern Route days in the Metropolitan subregion.

Exceedances are more frequent in subregions further inland from the coastal and metropolitan areas. With the sole exception of the Mountain subregion, the greatest frequency of exceedance days in these areas occurs under the Partial Southern Route Pattern. In the San Gabriel Valley, Inland Foothill and Inland Valley subregions, the Typical category also includes a relatively high frequency of exceedance days. The Inland Metropolitan subregion experiences its greatest frequency of exceedance days under the Partial Southern Route Pattern, and in the Mountain subregion, exceedances are most frequent on Partial Eddy days.

Pphm-Hours above 9 pphm

To obtain a more complete picture of daily ozone concentrations and their potential health effects than can be seen simply by analyzing the daily maximum concentration, we also calculated within-category distributions of pphm-hours per day above 9 pphm. These values were obtained as follows: For each day at each station, we summed the difference between the hourly average ozone concentrations and 9 pphm, counting only those hours in which this difference was positive. These individual station pphm-hour values were then averaged over the stations in each subregion to obtain the subregion average pphm-hours.

We compared results obtained through the above procedure using (1) all days, and (2) only days with valid daily maximum concentrations. No significant differences were observed between these two sets of results. Boxplots of the distributions over

TABLE 3-8. Number of days in each source history category on which subregion average daily maximum ozone concentration exceeds 20 pphm.

Subregion	Category								Missing
	1	2	3	4	5	6	7	8	
North Coast	0	0	0	0	0	1	0	0	0
South Coast	0	0	0	0	0	1	0	0	0
Metropolitan	0	0	0	0	1	1	0	0	0
San Fernando Valley	2	3	1	0	5	1	0	0	0
San Gabriel Valley	0	0	1	0	9	4	0	0	1
Inland Metropolitan	5	12	9	2	15	7	1	0	7
Inland Foothill	3	15	7	1	16	1	1	0	5
Inland Valley	5	14	6	0	11	2	0	0	4
Mountain	7	19	2	0	1	1	0	0	1

valid days are presented in Figure 3-35. Differences between categories within each subregion are generally the same as those described above for the daily maximum concentration.

DETERMINATION OF METEOROLOGICAL SIGNATURES

The objective procedure for ozone pattern classification discussed above makes use of the observed spatial distribution of relative ozone concentrations to infer the source-history characteristics of a particular day. This approach can be expected to yield consistent pattern classifications only so long as there are no major changes in the magnitude and geographical distribution of precursor emissions. Therefore, inconsistent results may be obtained if the objective criteria were to be applied to years far removed from the time period we examined (1983 - 1985; Tuesdays-Thursdays). To remedy this situation, it is necessary to develop a means of identifying ozone pattern classifications based solely on meteorological characteristics.

We attempted to identify sets of meteorological conditions that are uniquely associated with each of the ozone categories described above by using CART to grow classification trees in which the objectively determined categories are the dependent variable, and the daily meteorological variables are the predictor variables. Such a tree based on the meteorological variables in Table 3-9 for the period 1983 - 1985 is depicted in Figure 3-36. Table 3-10 shows the distribution of categories by terminal node. The first split in the tree is based on the wind direction at San Diego. This split serves to separate Eddy days (category #1) from all other days since wind directions on Eddy days have a southerly component. Other key variables used in the classification tree include T850 and MAXDIF, which serve to separate Low Ozone days (category #8) from other days; PGLAXWJF, which is negative on Partial Southern Route (category #5), Southern Route (category #6), and Offshore (category #7) days; and LATHPR07 and LATHPR16, which are near zero or slightly negative on offshore days but not on other days.

To use the classification tree described above as a decision rule, we assign to each terminal node the category represented by the majority of days in the node. These categories are identified by the lower numbers in the boxes in Figure 3-36. If a given day is run down the tree, the predicted category for the day is the category assigned to the terminal node the day falls into. If the entire sample of days used to construct the tree in Figure 3-36 is run back down the tree, 25 percent of these days will have predicted categories that differ from their actual categories as determined by the objective method. In other words, the relative resubstitution misclassification rate of the tree in Figure 3-36 is 25 percent. Since this misclassification rate is based on the actual sample of days used to construct the tree in the first place, it can be expected to be lower than the rate that would be obtained by running a large

TABLE 3-9. Meteorological variables used in CART analysis.

Name	Description
MAXTEMP	Daily maximum temperature at Thermal (deg. F).
TMAXSBD	Daily maximum temperature at San Bernardino (deg. F).
TMAXLAX	Daily maximum temperature at LAX (deg. F).
MAXDIFF1	TMAXSBD - TMAXLAX (deg. F).
MAXDIFF2	MAXTEMP - TMAXLAX (deg. F).
T850	850 mb temperature from 1300 UTC UCLA/LMU sounding (deg. C).
DELTAT	Temperature difference across inversion (Top - Base; deg. C).
DELTAPG	24-hour change in PGLAXWJF (mb).
WS0700	San Diego wind speed, 7:00 a.m. LST (kts).
WS1000	San Diego wind speed, 10:00 a.m. LST (kts).
LATHPR07	Pressure gradient (LAX - Thermal), 7:00 a.m. LST (mb).
LATHPR16	Pressure gradient (LAX - Thermal), 4:00 p.m. LST (mb).
BASEHT	Inversion base height at UCLA/LMU, 1300 UTC (feet).
PGLAXWJF	Pressure gradient (LAX - Lancaster), 1300 UTC (mb).
AWSLGB	Daily scalar average wind speed, Long Beach.
WD0700	Wind direction, San Diego, 7:00 a.m. LST (1 = N, 2 = NE, ..., 8 = NW).
WD1000	Wind direction, San Diego, 10:00 a.m. LST (1 = N, 2 = NE, ..., 8 = NW).
RWDLGB	24-hour resultant wind direction, Long Beach (1 = N, 2 = NE, ..., 8 = NW).

TABLE 3-10. Distribution of source history categories by terminal node for the classification tree depicted in Figure 3-36.

Terminal Node	Category							
	1	2	3	4	5	6	7	8
1	0	0	0	0	0	0	6	1
2	0	1	0	0	0	10	0	2
3	0	3	0	0	20	8	0	0
4	0	0	0	0	0	0	0	8
5	0	44	18	0	6	0	0	2
6	1	0	0	0	0	0	1	20
7	47	0	0	6	0	0	0	3

independently selected sample of days down the tree. An estimate of this misclassification rate is calculated by CART using a cross-validation procedure. For the tree shown in Figure 3-36, this cross-validation misclassification rate is 31 percent.

Although the classification tree just described does a reasonably good job of identifying the key meteorological aspects of each ozone category, the tree failed to identify any significant new relationships between weather conditions and ozone patterns that were not already accounted for in the objective classification procedure (one minor exception is the identification of splits on T850 and MAXDIF1 that can be used to characterize Low Ozone days). Thus, by using just the available meteorological data without reference to relative ozone concentrations, it is not possible to determine the source history category to which a day is best assigned without a fair degree of uncertainty. This makes it difficult to perform a consistent categorization of days for years falling outside the 1983-1985 Tuesday-Thursday time frame. It is possible that the inclusion of additional meteorological data in the analysis would help to alleviate this problem.

OZONE - METEOROLOGY RELATIONSHIPS WITHIN PATTERNS

Each of the ozone patterns described above is primarily intended to represent a particular flow or source history pattern. Thus up to this point, we have been principally concerned with those meteorological aspects of each pattern that relate to air parcel trajectories. As a result, other meteorological characteristics that are more closely related to a day's ozone formation potential are generally not affected by the categorization process. Therefore, the days in a single category include a wide range of meteorological factors that relate to ozone formation potential. For example, two days within a single category may have fairly different 850 mb temperatures and thus fairly different maximum ozone concentrations. Table 3-11 lists those meteorological variables in our database that we believe to be more closely associated with ozone formation potential than with flow pattern determinations.

For the purpose of distinguishing between days in the same source history category that may have significantly different ozone formation potentials and thus may play different roles in the development of emission control scenarios, we examined the relationships between various meteorological factors and ozone concentrations within each source history category. A step-wise linear regression analysis was carried out in which the subregion average daily maximum ozone concentrations served as the dependent variable, and the meteorological variables in Table 3-11 served as the predictor variables. A separate regression was performed for each subregion for the Eddy and Typical Pattern source history categories. These categories were selected because they are associated with clearly different meteorological regimes and contain a reasonably large number of days on which to base the analysis. Results are summarized in Table 3-12. Only those meteorological variables with coefficients significantly different from zero at the 85 percent confidence level are included in the regression equations for each subregion/pattern

TABLE 3-11. Meteorological variables used in regression analysis.

Name	Description
LATHPR07	Pressure gradient (LAX - Thermal), 7:00 a.m. LST (mb).
LATHPR16	Pressure gradient (LAX - Thermal), 4:00 p.m. LST (mb).
PGLAXWJF	Pressure gradient (LAX - Lancaster), 1300 UTC ¹ (mb).
AWSLGB	Daily scalar average wind speed, Long Beach.
TMAXSBD	Daily maximum temperature at San Bernardino (deg. F).
DELTAT	Temperature difference across inversion (Top - Base; deg. C).
T850	850 mb temperature from 1300 UTC UCLA/LMU sounding (deg. C).
DELTAPG	24-hour change in PGLAXWJF (mb).
TMAXLAX	Daily maximum temperature at LAX (deg. F).
BASEHT	Inversion base height at UCLA/LMU, 1300 UTC (feet).
MAXDIFF1	TMAXSBD - TMAXLAX (deg. F).
MAXDIFF2	MAXTEMP - TMAXLAX (deg. F).
LAGT850	Previous day's T850 (°C).

¹ Universal Time Coordinates, i.e., local time at the prime meridian or Greenwich Mean Time.

TABLE 3-12. Stepwise regression results. Meteorological variables are defined in Table 3-11. MAXCON vs meteorological variables by TESPATT for patterns 1 and 3 by subregions.

Subregion	Category	Met. Variable	Coefficient	R ²	
				Partial	Model
1	1	TMAXSBD	0.16	0.11	
		LAGT850	-0.16	0.16	0.27
1	3	DELTAT	0.47	0.35	
		PGLAXWJF	-1.20	0.21	0.56
2	1	LAGT850	-0.10	0.22	0.22
2	3	PGLAXWJF	-1.00	0.23	0.23
3	1	TMAXSBD	0.18	0.20	
		LAGT850	-0.12	0.08	0.28
3	3	DELTAT	0.65	0.55	
		TMAXSBD	0.16	0.10	0.66
4	1	TMAXSBD	0.27	0.44	
		LAGT850	-0.20	0.04	0.48
		T850	0.26	0.03	0.50
4	3	DELTAT	1.14	0.40	0.40
5	1	TMAXSBD	0.51	0.43	
		TMAXLAX	-0.48	0.05	0.47
		LATHPRO7	-0.74	0.07	0.54
5	3	DELTAT	1.39	0.61	
		TMAXSBD	0.38	0.14	0.75
6	1	----- none -----			
6	3	DELTAT	1.34	0.52	0.52
7	1	TMAXSBD	0.38	0.41	
		TMAXLAX	-0.54	0.10	0.50
		LATHPRO7	-0.48	0.06	0.56
		T850	0.27	0.03	0.59

continued

TABLE 3-12. concluded

Subregion	Category	Met. Variable	Coefficient	R ²	
				Partial	Model
7	3	DELTAT	1.27	0.44	
		TMAXSBD	0.32	0 09	0.53
8	1	T850	0.78	0.39	
		TMAXLAX	-0.51	0.09	0.49
		TMAXSBD	0.28	0.07	0.56
		LAGT850	-0.20	0.06	0.62
		DELTAT	-0.40	0.03	0.65
8	3	DELTAT	1.53	0.56	0.56
9	1	T850	0.62	0.49	
		LAGT850	-0.23	0.06	0.55
		TMAXSBD	0.33	0.04	0.59
		TMAXLAX	-0.30	0.04	0.64
9	3	DELTAT	1.17	0.67	0.67

combination.* Variables are listed in the order in which they entered into the regression, along with the increase in the fraction of variance explained resulting from their inclusion (Partial R^2). Model R^2 s listed in the table are equal to the sum of the partial R^2 s and indicate the total fraction of variance explained by the complete regression equation.

Examination of the results in Table 3-12 reveals that only weak connections, as indicated by relatively low R^2 values, exist between meteorological conditions and ozone concentrations in the coastal and metropolitan subregions on eddy days (category #1). Surprisingly, none of the meteorological variables was found to have a significant correlation with ozone concentrations on eddy days in the Inland Metro subregion. As noted previously, these subregions experience low relative and absolute ozone concentrations on such days, and the results in Table 3-12 suggest that these concentrations are not strongly affected by the meteorological variability within this source history category. Recall that 39-52 percent of the variance in the coastal and metropolitan subregions is already accounted for by the source history classification. R^2 values for the Eddy Pattern in more inland subregions are all above 0.5. A low R^2 value was also obtained for the South Coast subregion on typical days (category #3), but in other subregions, R^2 values for Typical Pattern days are above 0.5 (except 0.40 in the San Fernando Valley).

On Typical Pattern days, the meteorological variable most closely associated with ozone concentrations is DELTAT, a measure of the inversion strength. Positive coefficients were calculated for this variable, indicating that ozone concentrations tend to be higher on days with stronger inversions. Significant positive coefficients were also obtained for TMAXSBD in the Typical Pattern day regressions in some subregions. For the coastal subregions, the onshore pressure gradient (PGLAXWJF) is also associated with the amount of ozone formed on Typical days since smaller positive gradients mean lighter winds and less influx of relatively clean ocean air.

On Eddy days, the inversion strength as measured by DELTAT does not bear a significant positive correlation to ozone levels. In fact, an inverse relationship seems to exist with this variable in the Inland Valley subregion. On the other hand, TMAXSBD is positively correlated with maximum ozone concentrations in most of the subregions on eddy days. Significant positive correlations were also obtained with T850 in some subregions. Surprisingly, an inverse relationship is indicated between the eddy day ozone concentrations and the previous day's 850 mb temperature (LAGT850) for six of the subregions. This may be a result of the fact that eddy days tend to occur early in the ozone season and thus are more frequently preceded by cooler days.

* An 85 percent confidence level was chosen so that all variables, including those that are only weakly related to ozone concentrations, would be considered by the step-wise regression procedure.

Cooler days along the coast (as measured by TMAXLAX) are associated with higher ozone concentrations on eddy days in the Inland Metro, Inland Foothill, Inland Valley, and Mountain subregions, as evidenced by the negative coefficients for this variable. This follows from the fact that sufficiently vigorous onshore winds are needed to transport ozone to these areas, and such winds result in cooler temperatures along the coast.

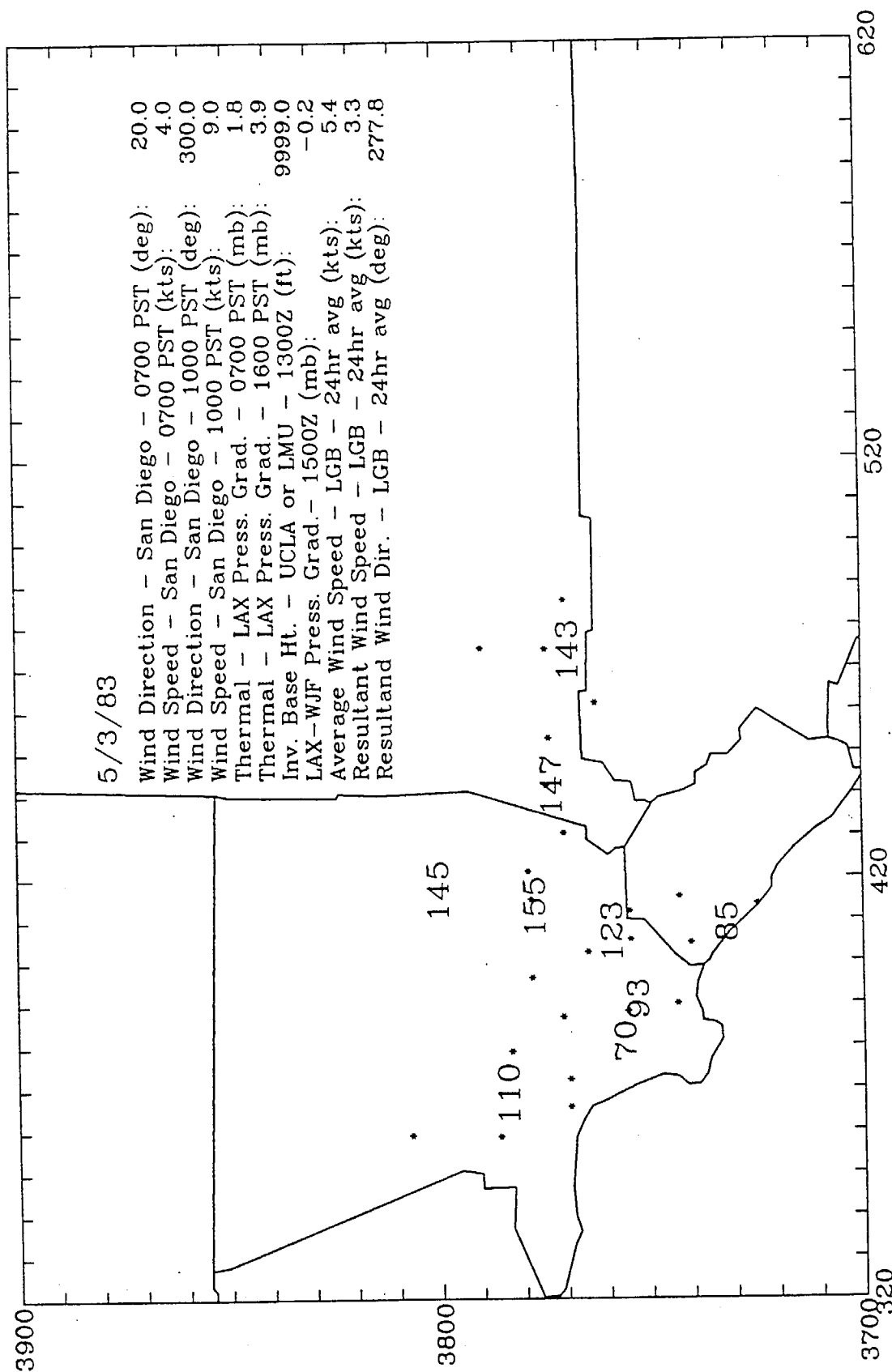


FIGURE 3-1. Example data map for 5/3/83. Large numbers superimposed on map indicate subregion average daily maximum ozone concentrations; (*) indicate station locations (see key to subregions, Figure 2-1).

Box plot for variable RATNET1, *==mean
 order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

(a) North Coast

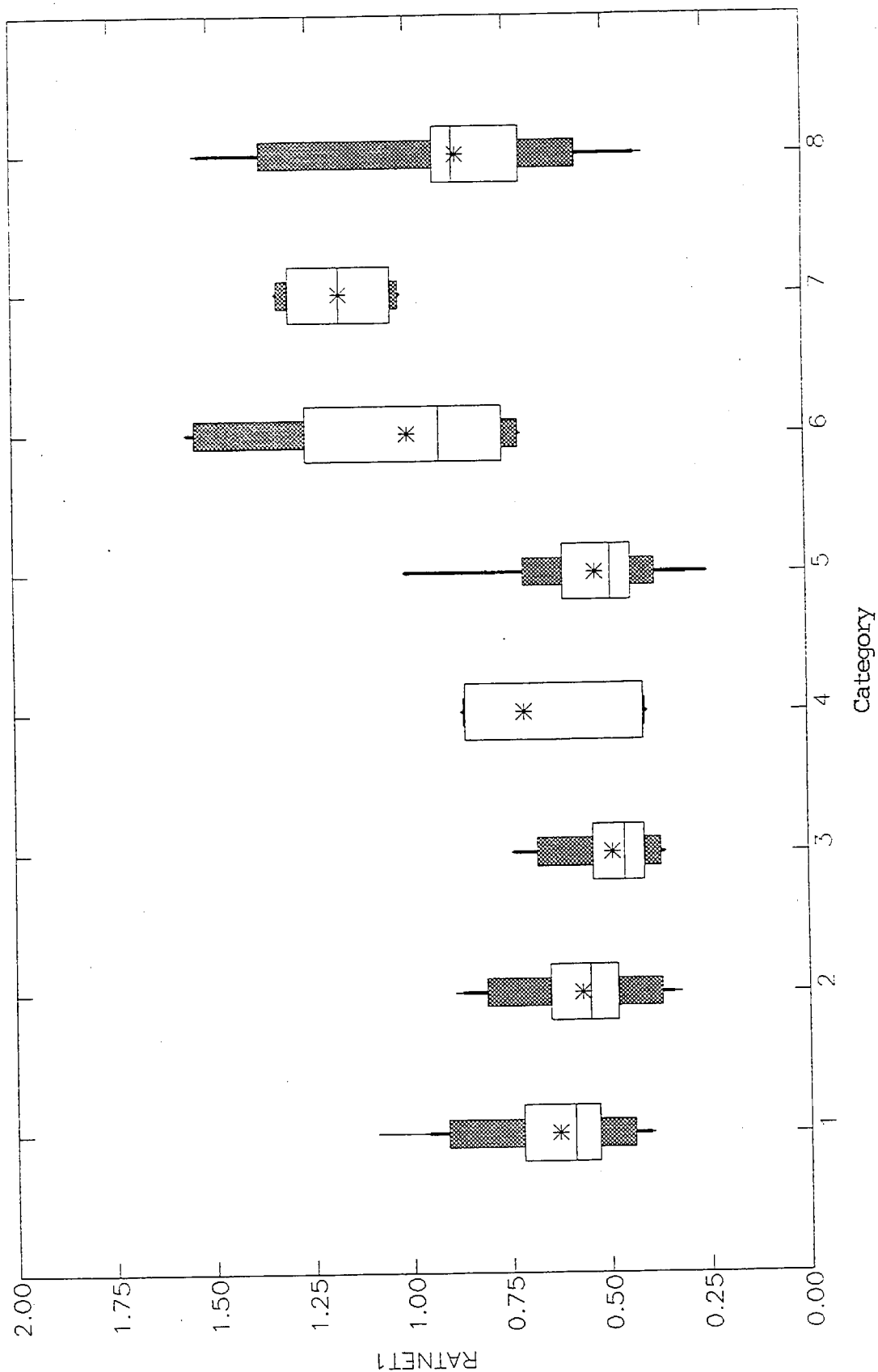


FIGURE 3-2. Box plots of relative subregion average daily maximum ozone concentrations for each source history category. (Key to box plots appears in Figure 3-3.)

Box plot for variable RATNET2, *=mean
 order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

(b) South Coast

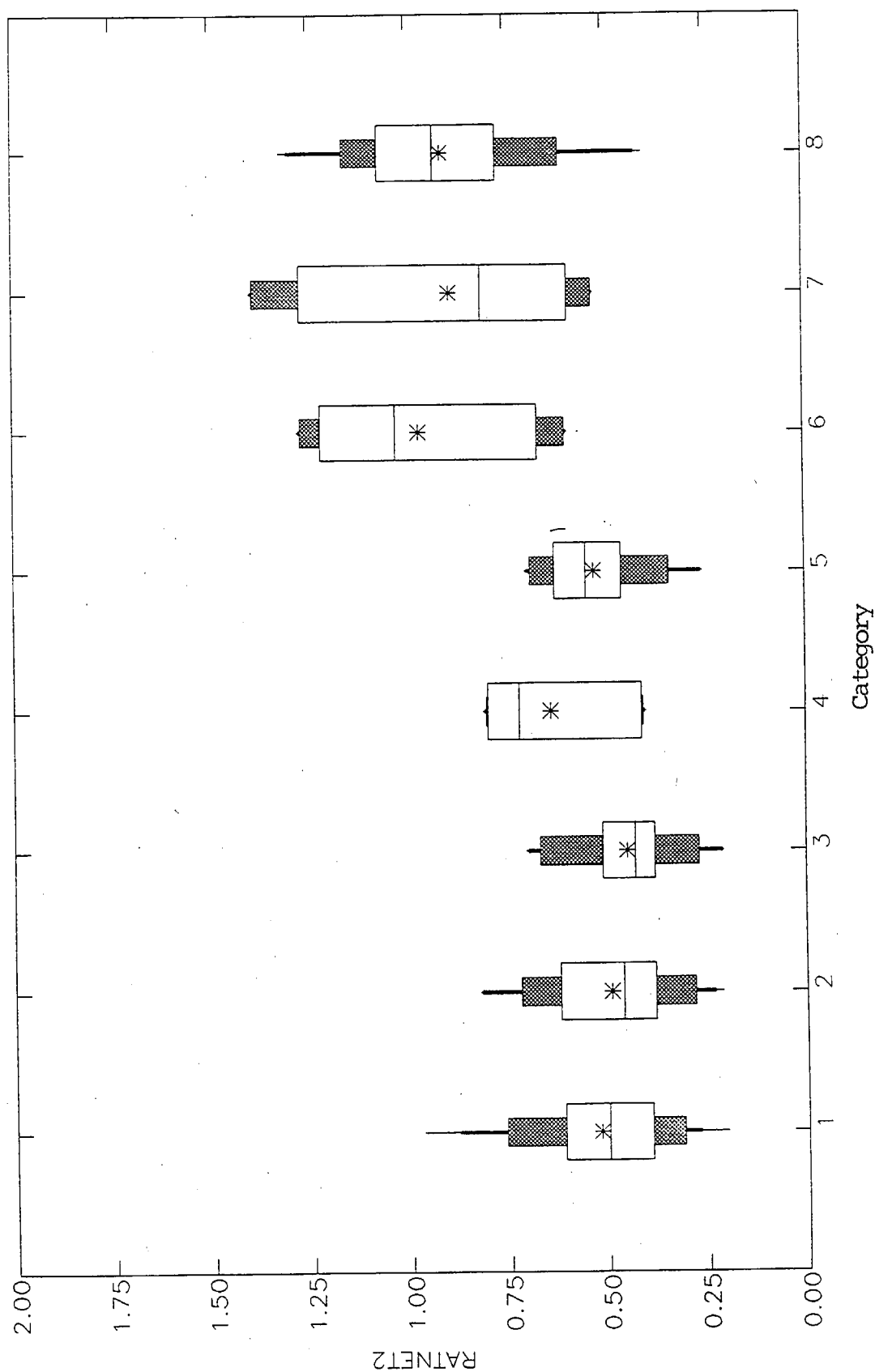


FIGURE 3-2. Continued.

Box plot for variable RATNET3, *==mean
order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

(c) Metropolitan

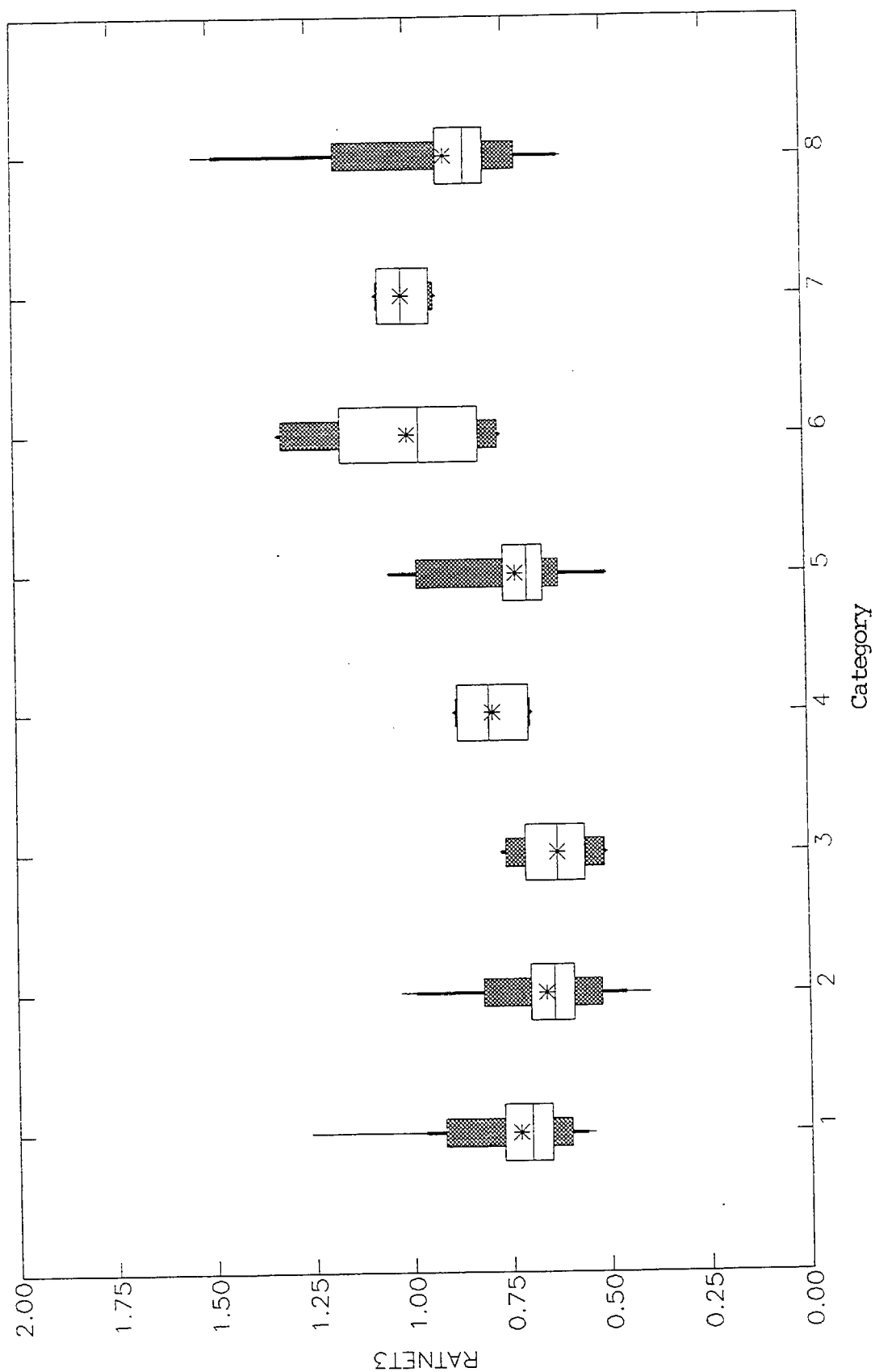


FIGURE 3-2. Continued.

Box plot for variable RAINET14, * = mean
order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

(d) San Fernando Valley

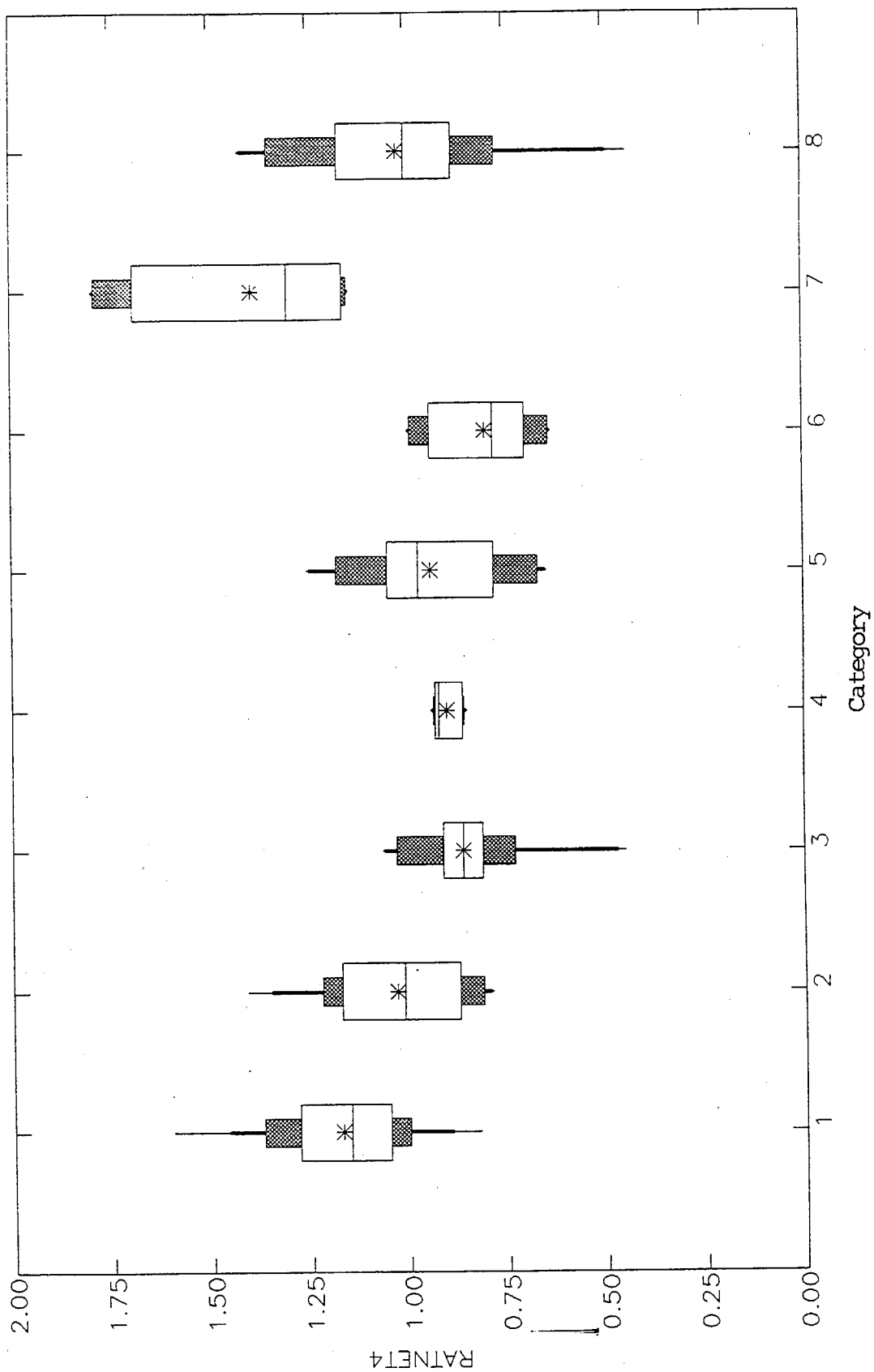


FIGURE 3-2. Continued.

Box plot for variable RATNET5, * = mean
order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

(e) San Gabriel Valley

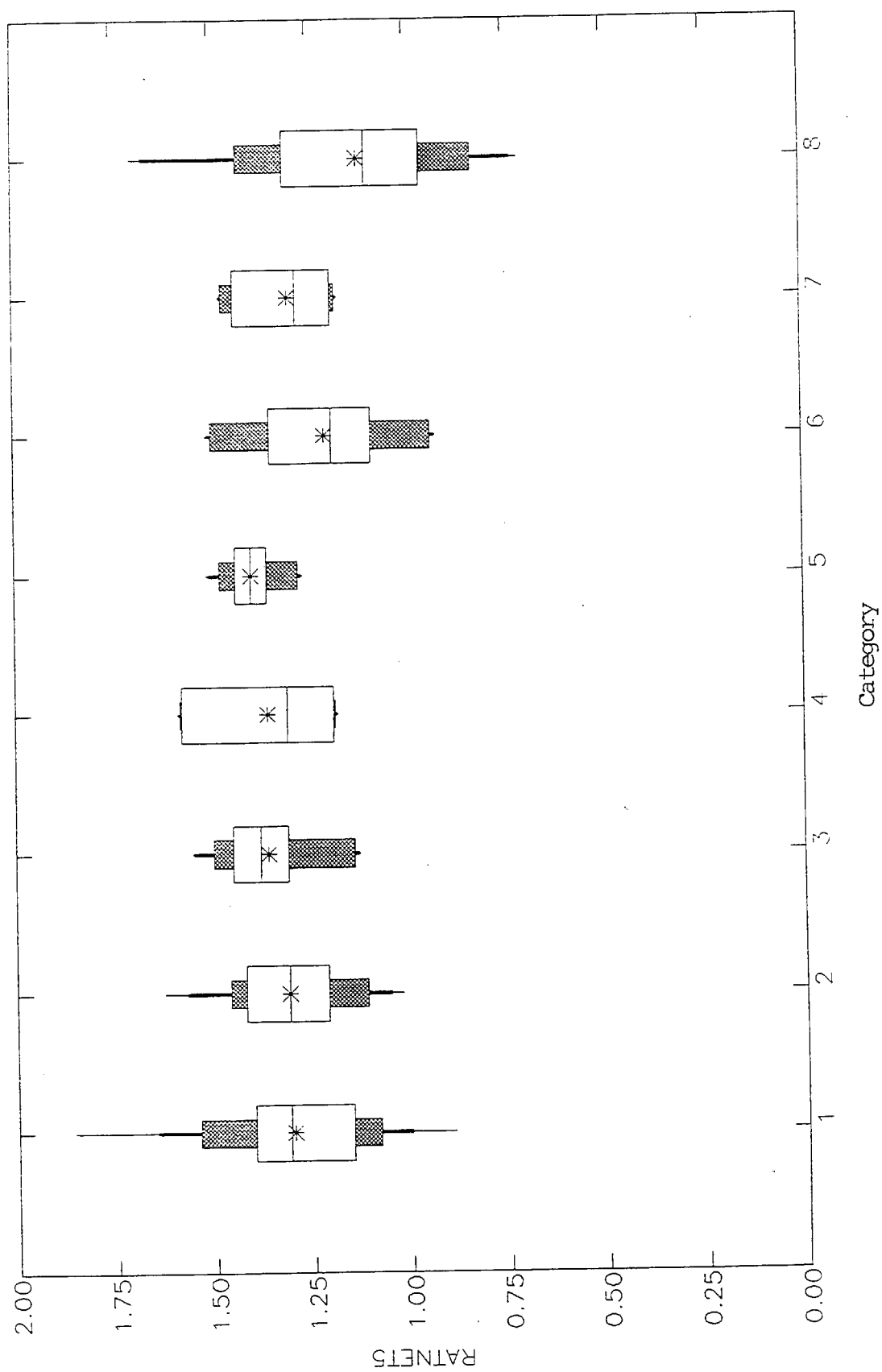


FIGURE 3-2. Continued.

Box plot for variable RAINET6, * = mean
 order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

(f) Inland Metropolitan

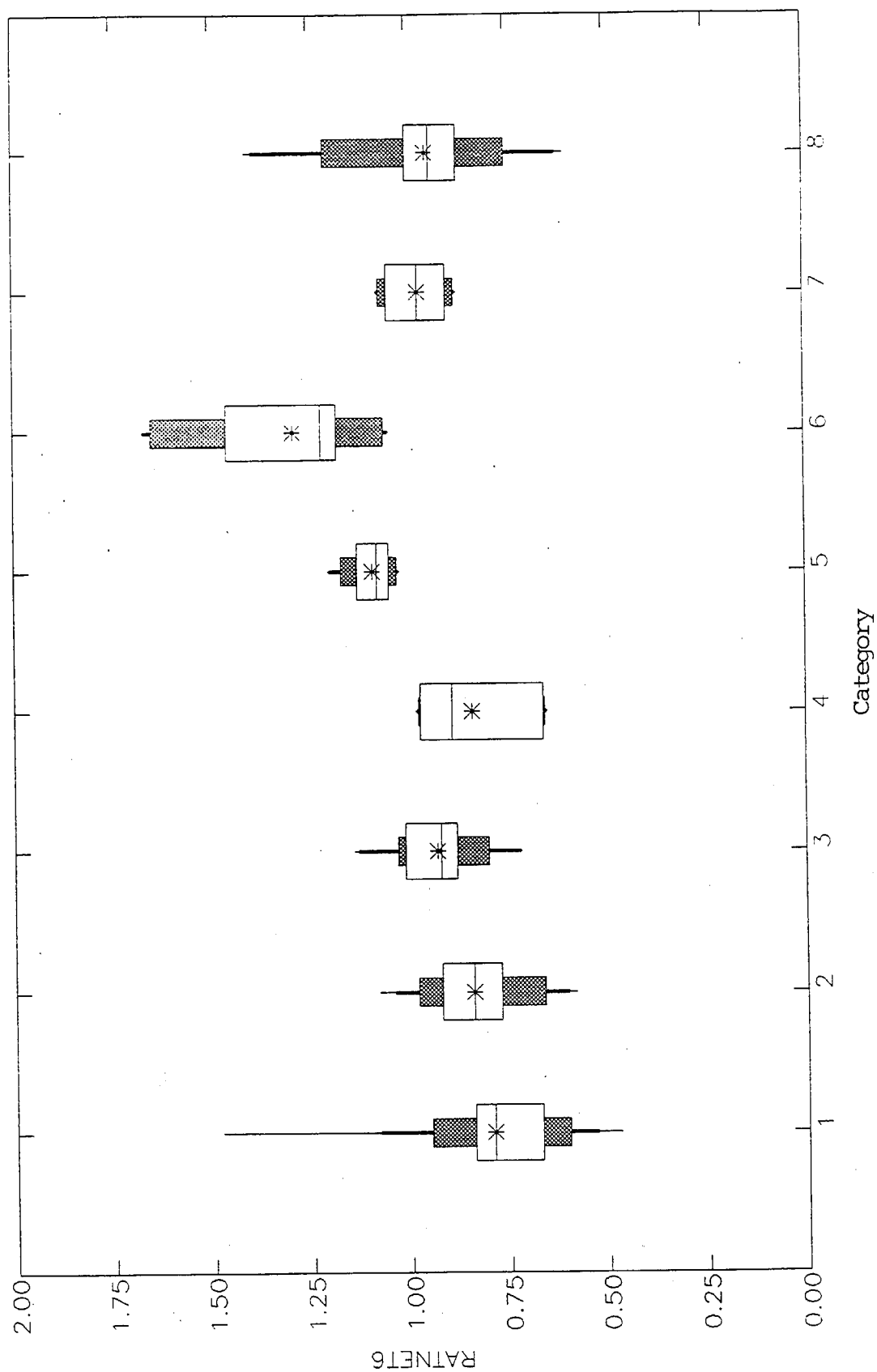


FIGURE 3-2. Continued.

Box plot for variable RATNET7, *=mean
order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

(g) Inland Foothill

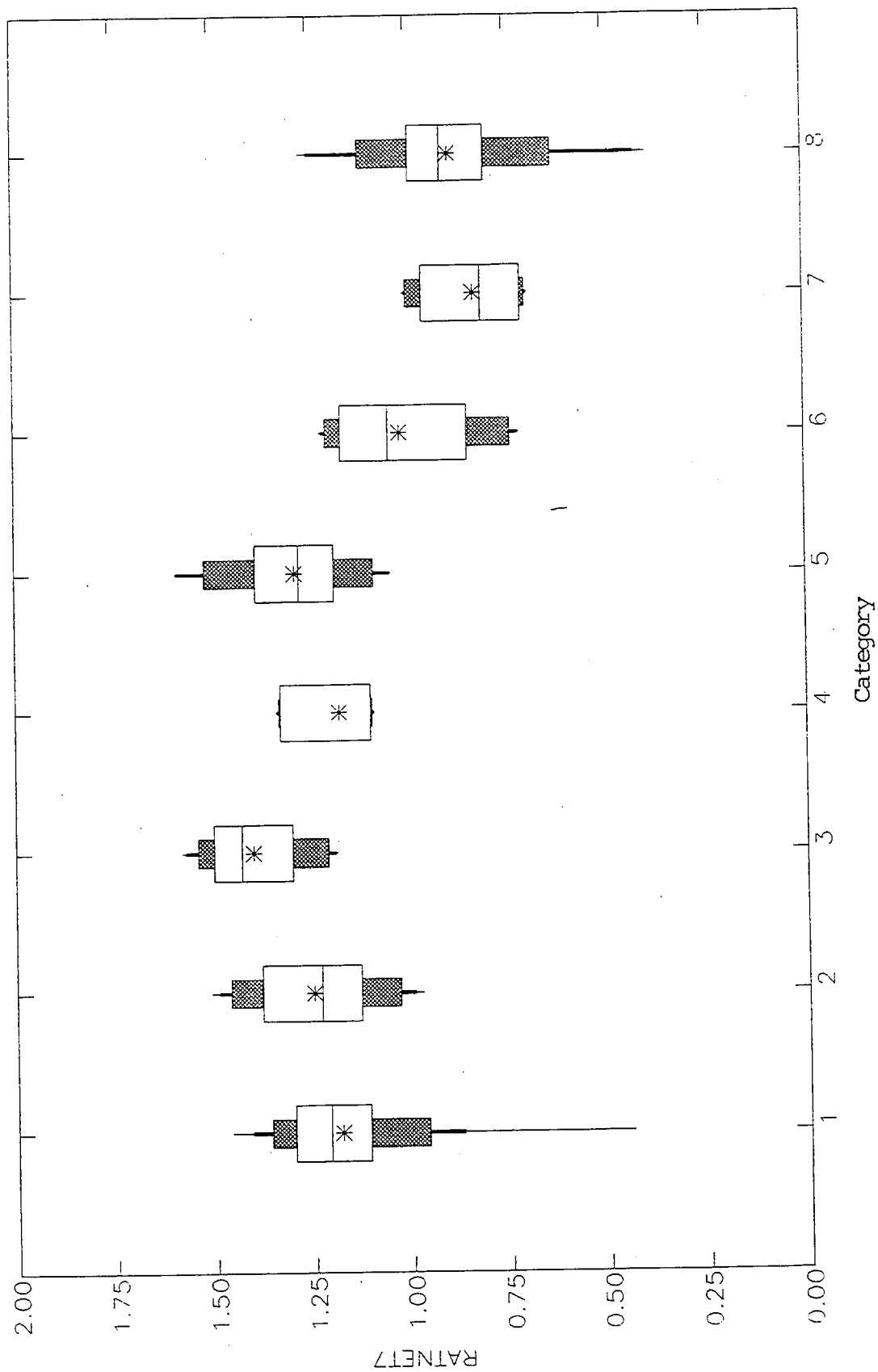


FIGURE 3-2. Continued.

Box plot for variable RAINET18, *==mean
 order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

(h) Inland Valley

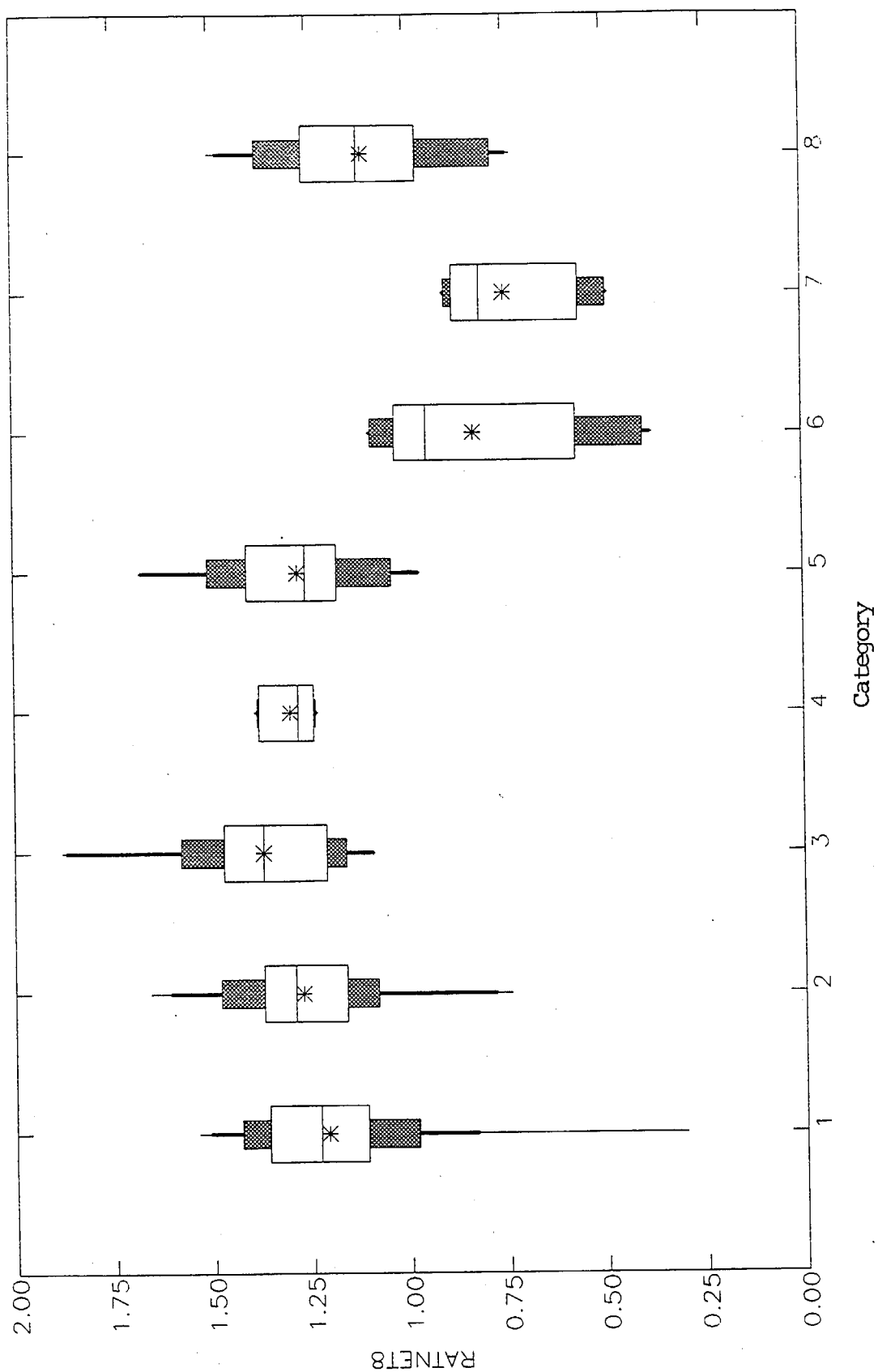


FIGURE 3-2. Continued.

Box plot for variable RATNET9, * = mean
 order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

(i) Mountain

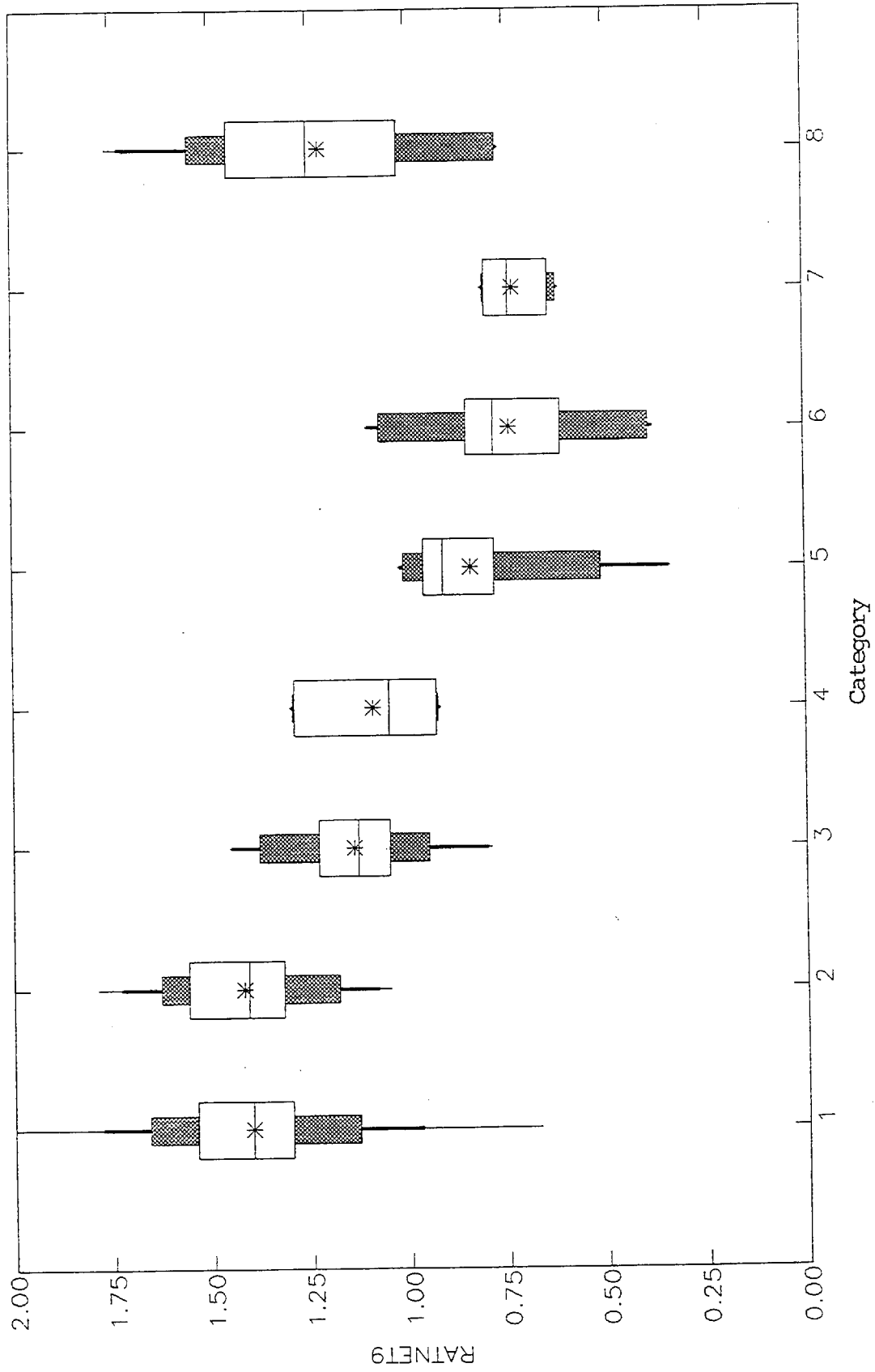


FIGURE 3-2. Concluded.

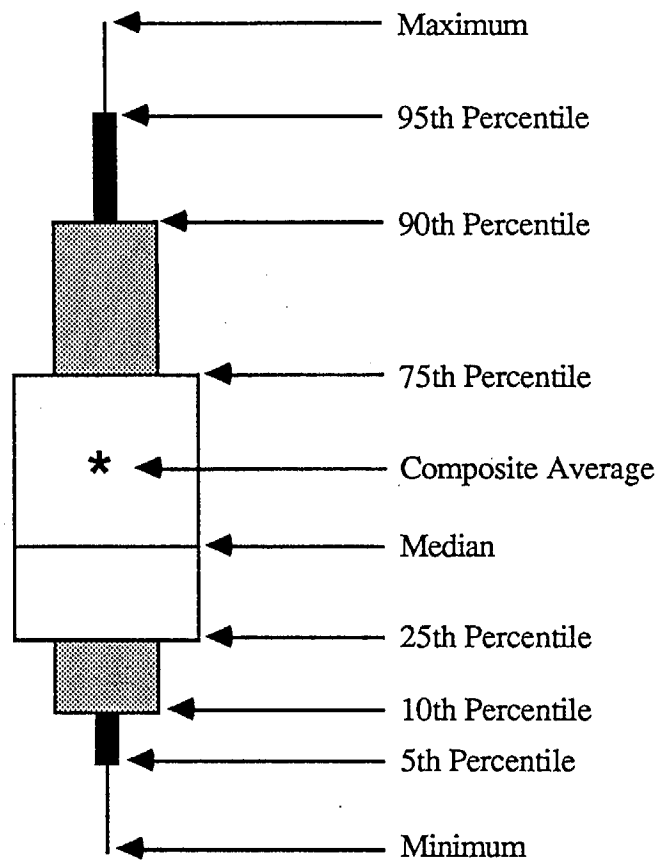


FIGURE 3-3. Key for boxplots in Figures 3-2 and 3-4 through 3-22.

(a) Eddy Pattern

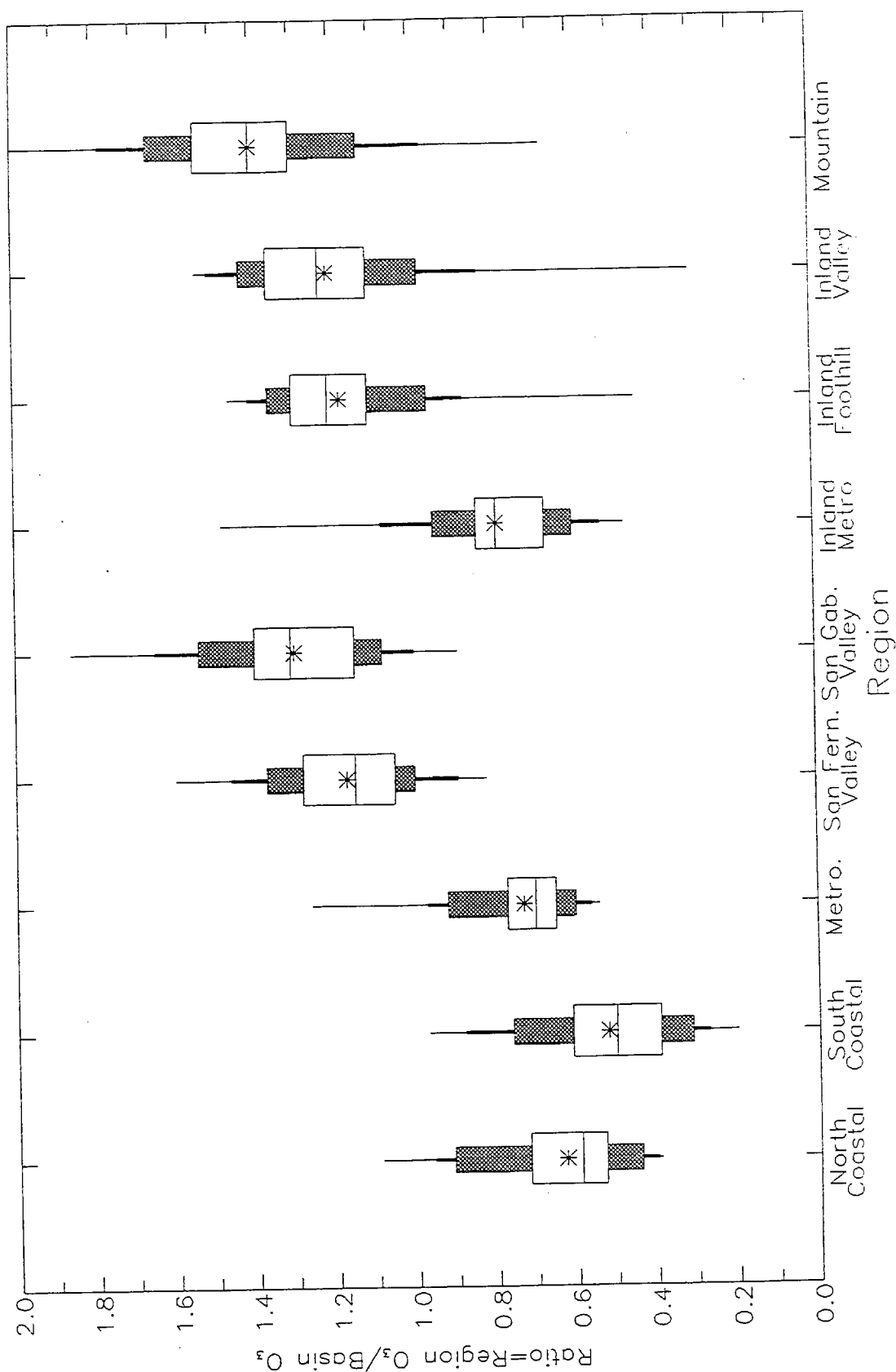


FIGURE 3-4. Box plots of relative subregion average daily maximum ozone concentrations. (Key to box plots is presented in Figure 3-3.)

(b) Partial Eddy Pattern

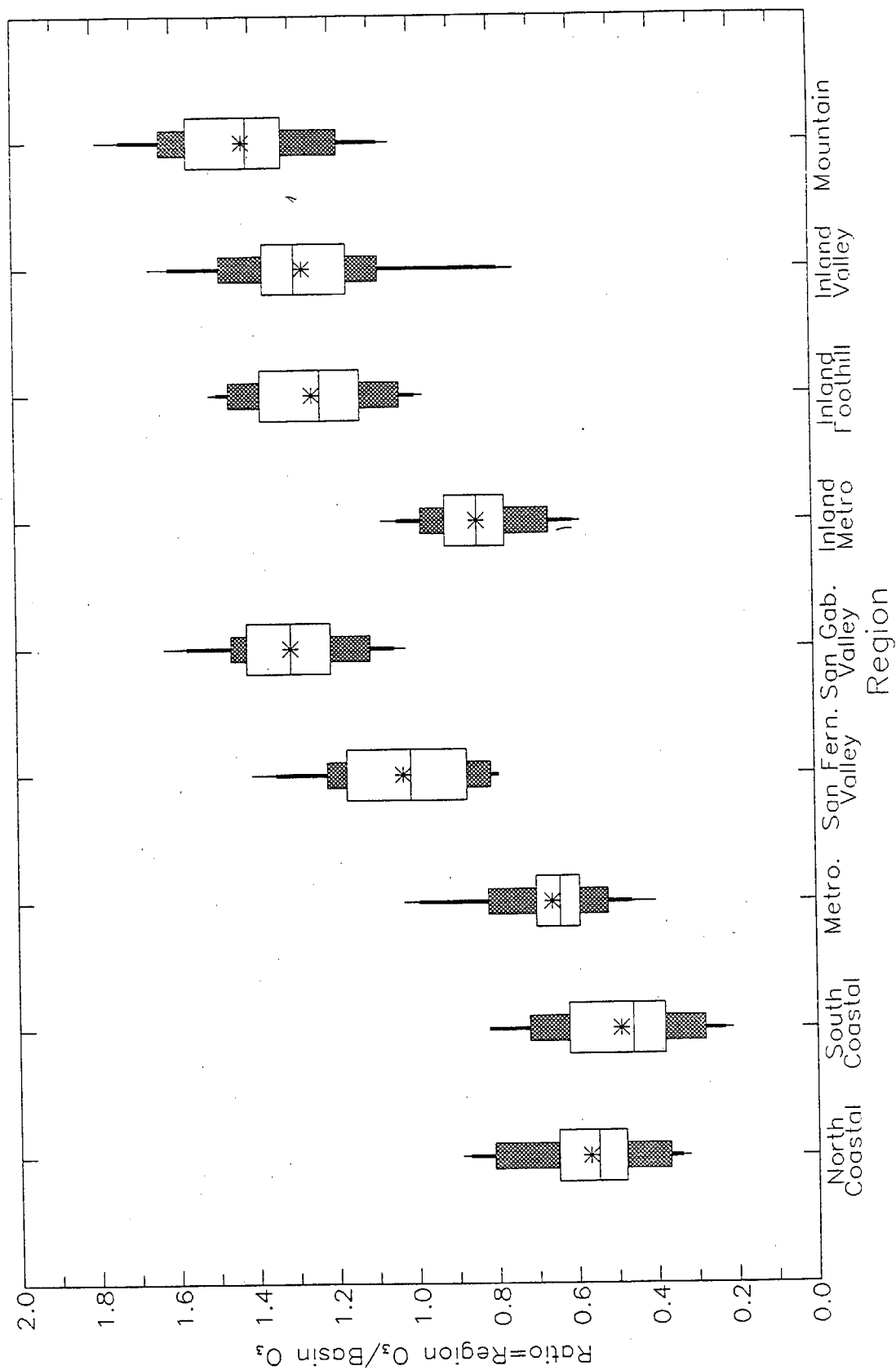


FIGURE 3-4. Continued.

(c) Typical Pattern

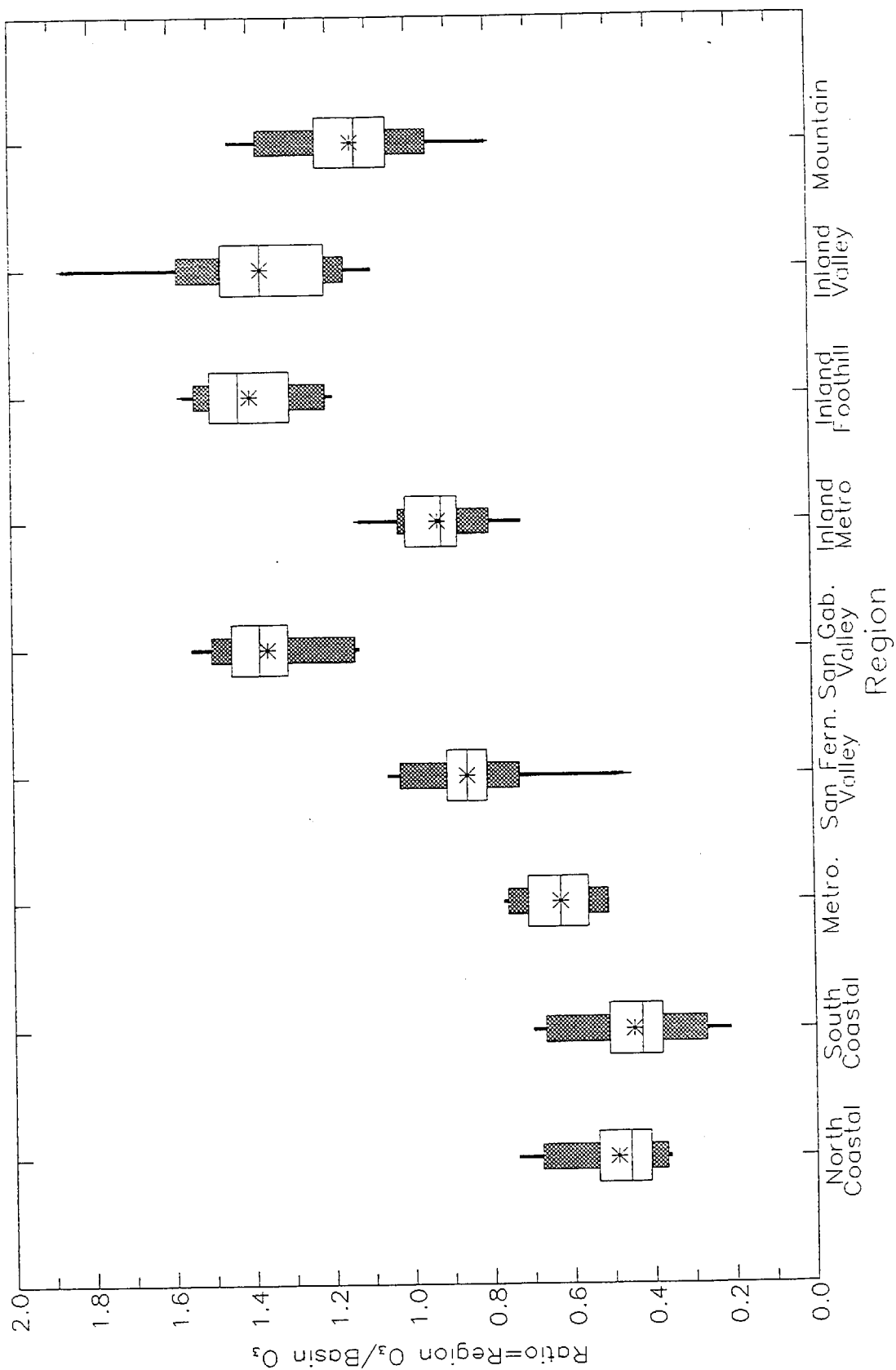


FIGURE 3-4. Continued.

(d) Typical Pattern - Eddy Winds

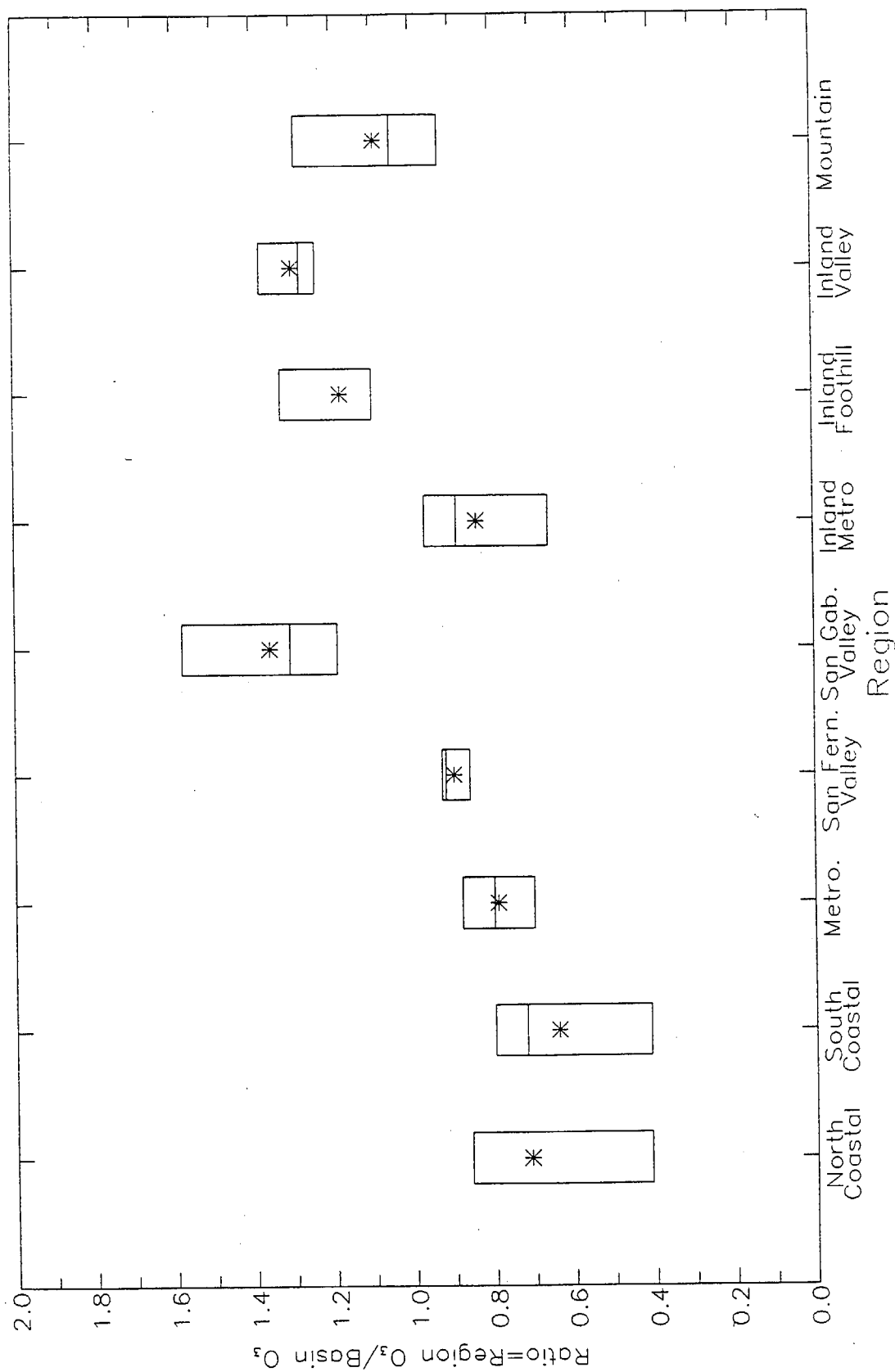


FIGURE 3-4. Continued.

(e) Partial Southern Route Pattern

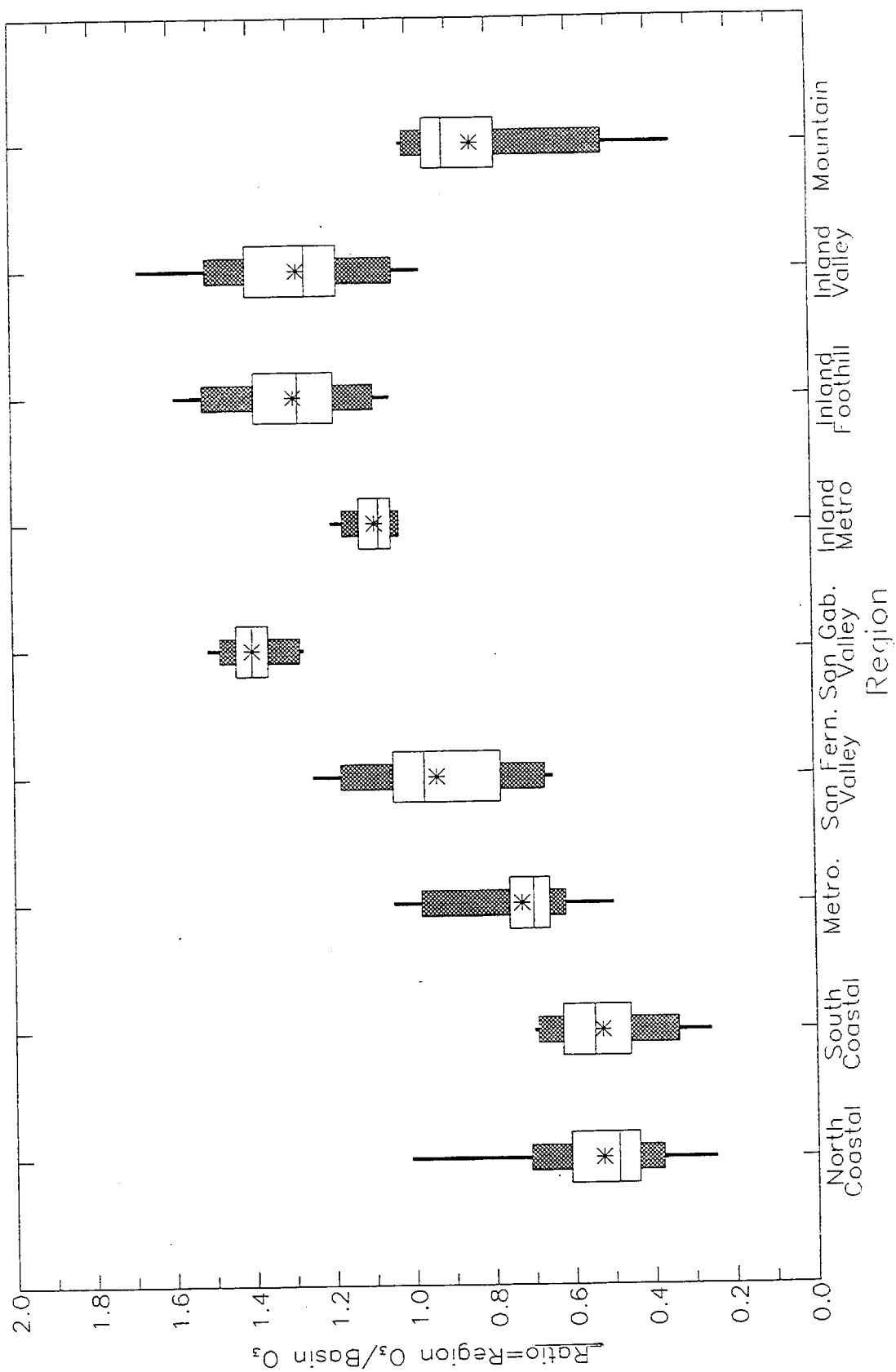


FIGURE 3-4. Continued.

(f) Southern Route Pattern

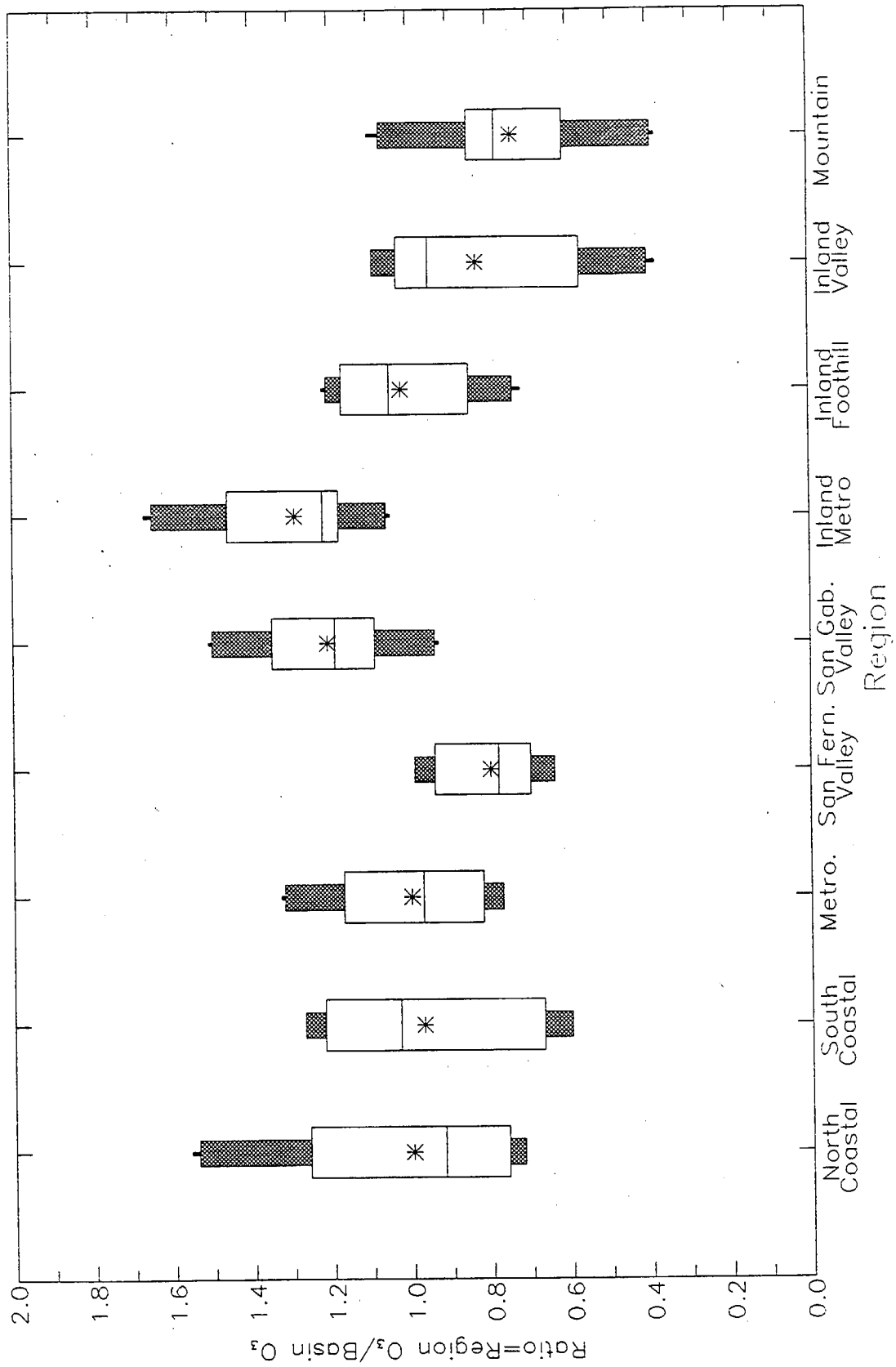


FIGURE 3-4. Continued.

(g) Offshore Pattern

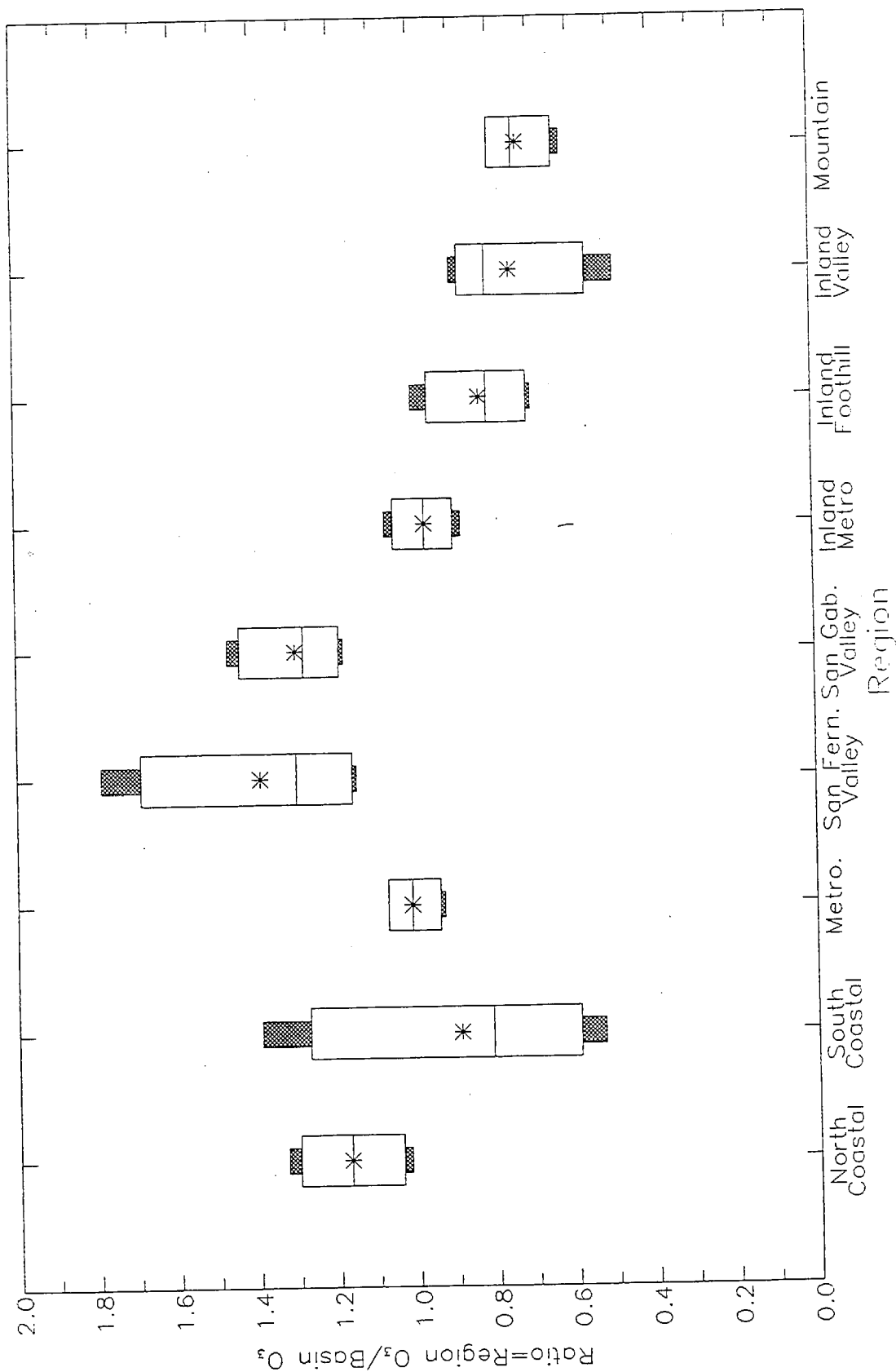


FIGURE 3-4. Continued.

(h) Low Ozone Pattern

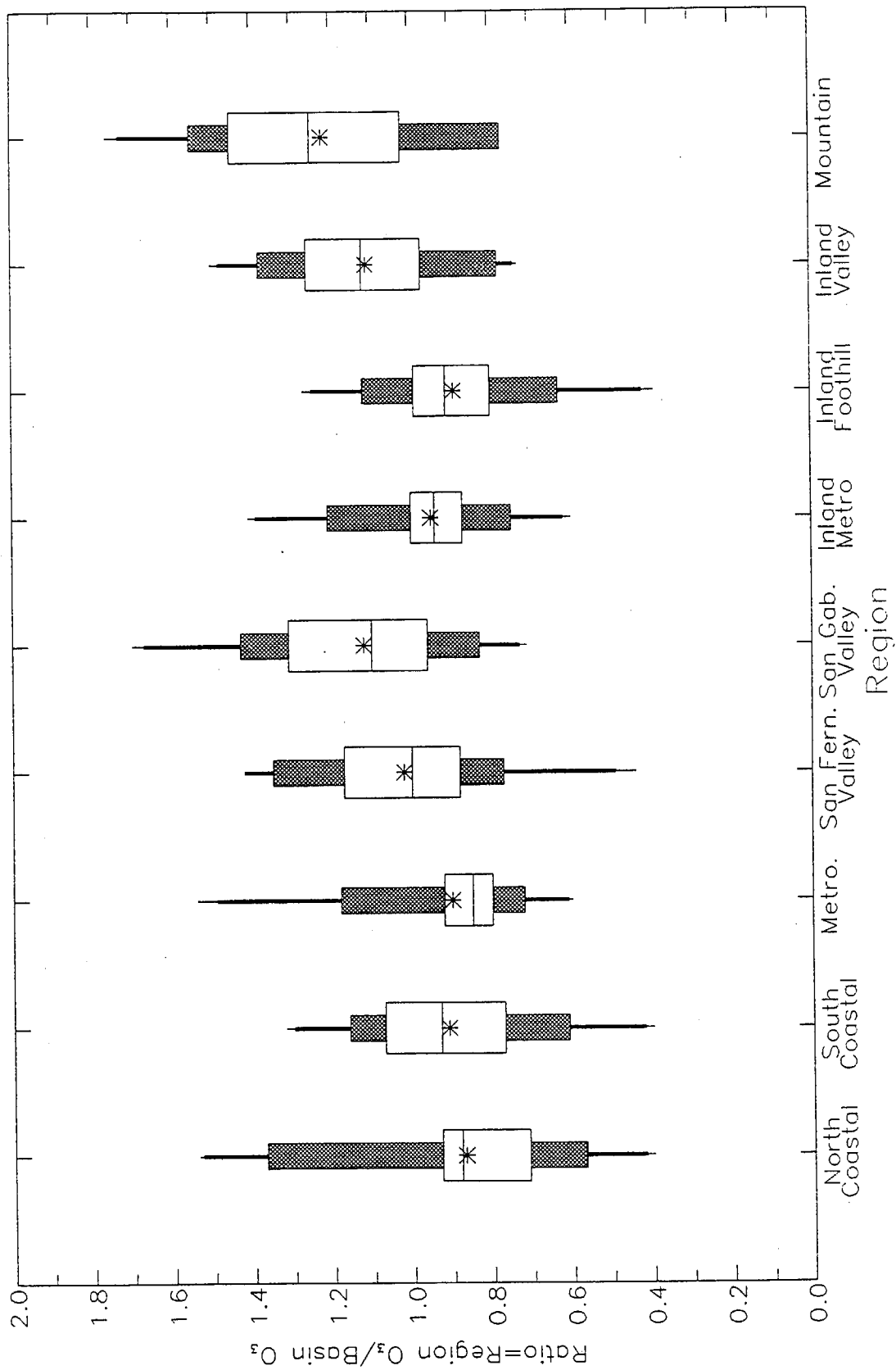


FIGURE 3-4. Continued.

(i) All Patterns

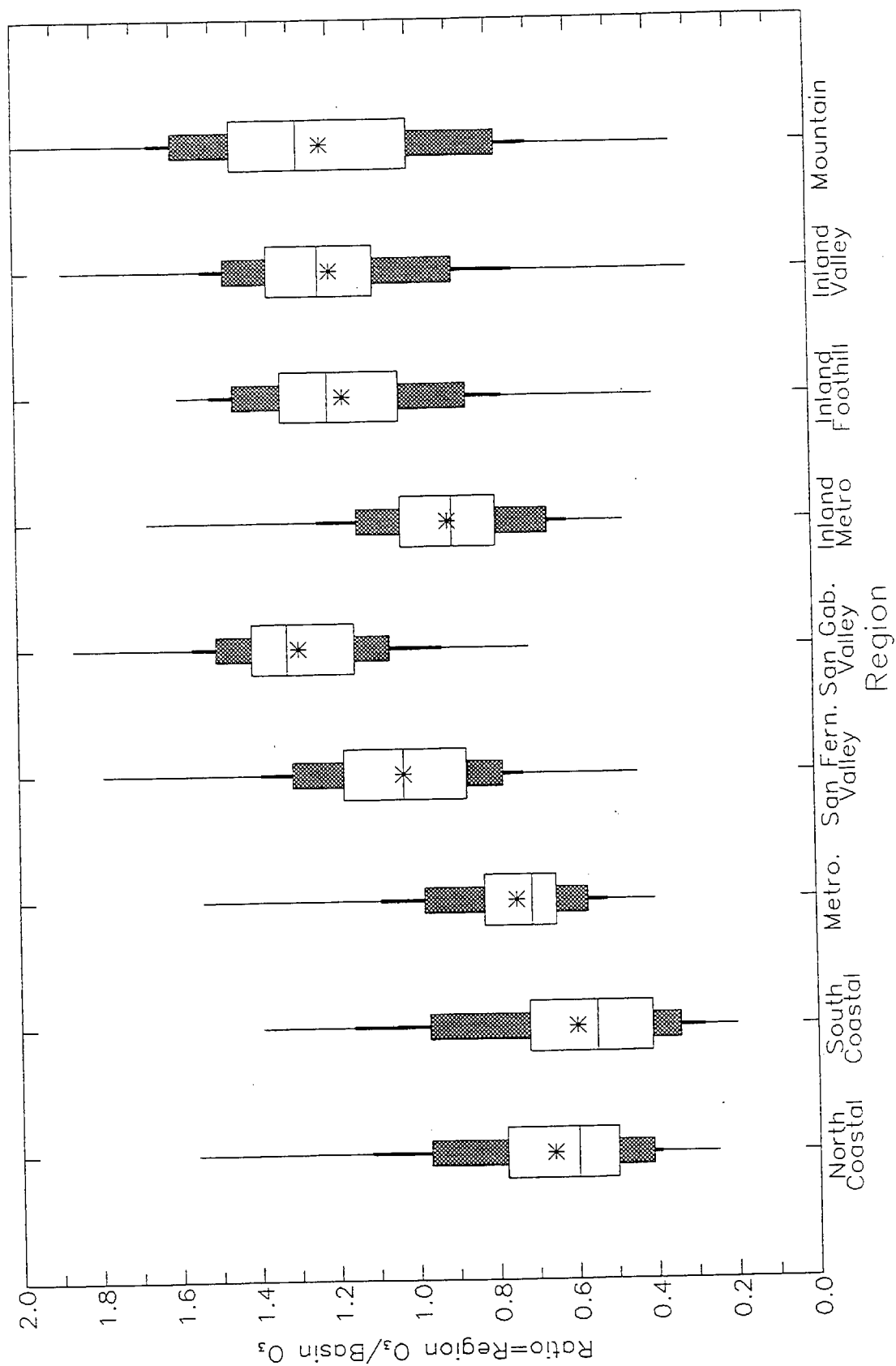


FIGURE 3-4. Concluded.

Box plot for variable MAXTEMP, *=mean
 order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

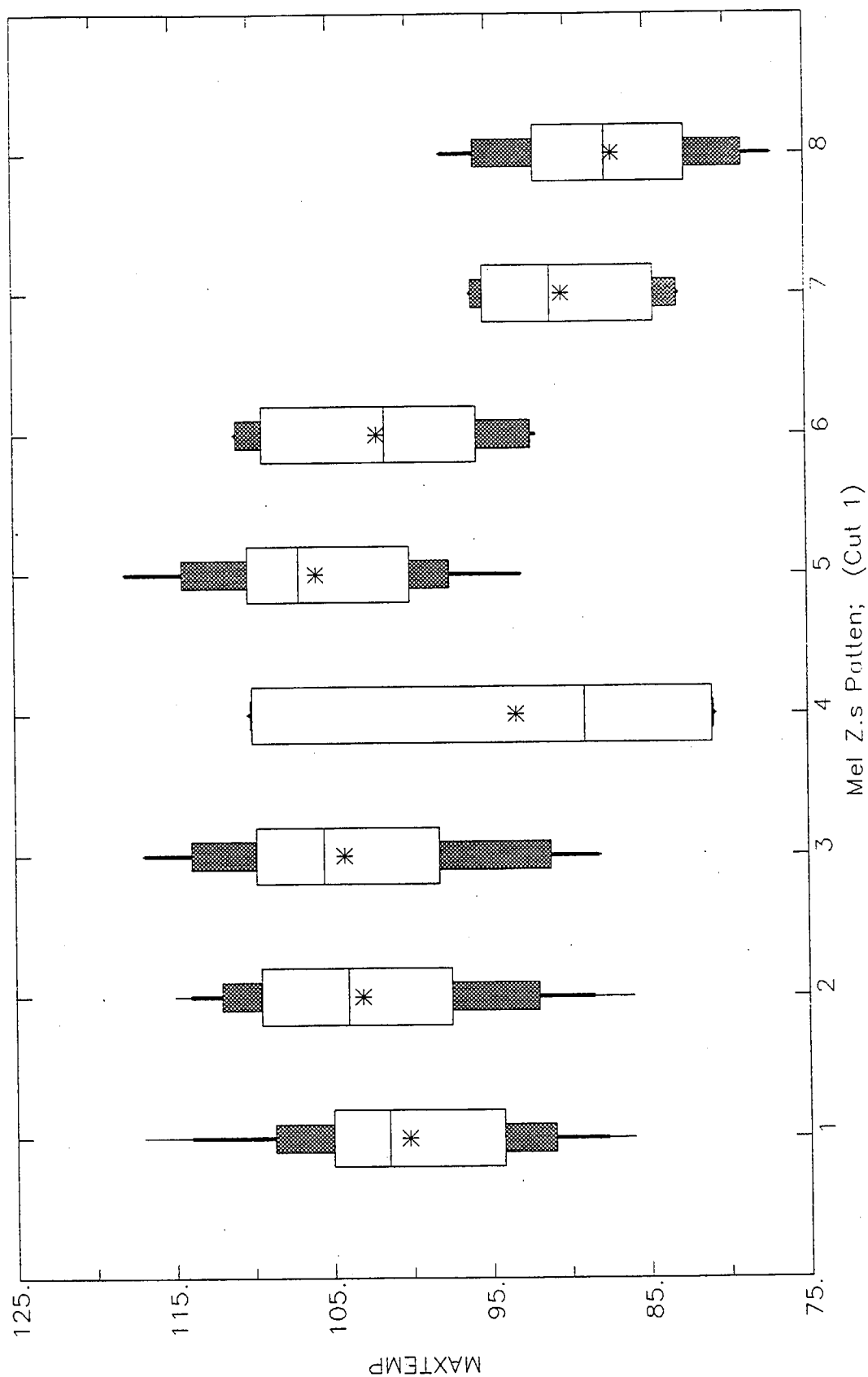


FIGURE 3-5. Boxplot of variable MAXTEMP for each source history category. (Key to boxplots appears in Figure 3-3.)

Box plot for variable WS0700, * = mean
order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

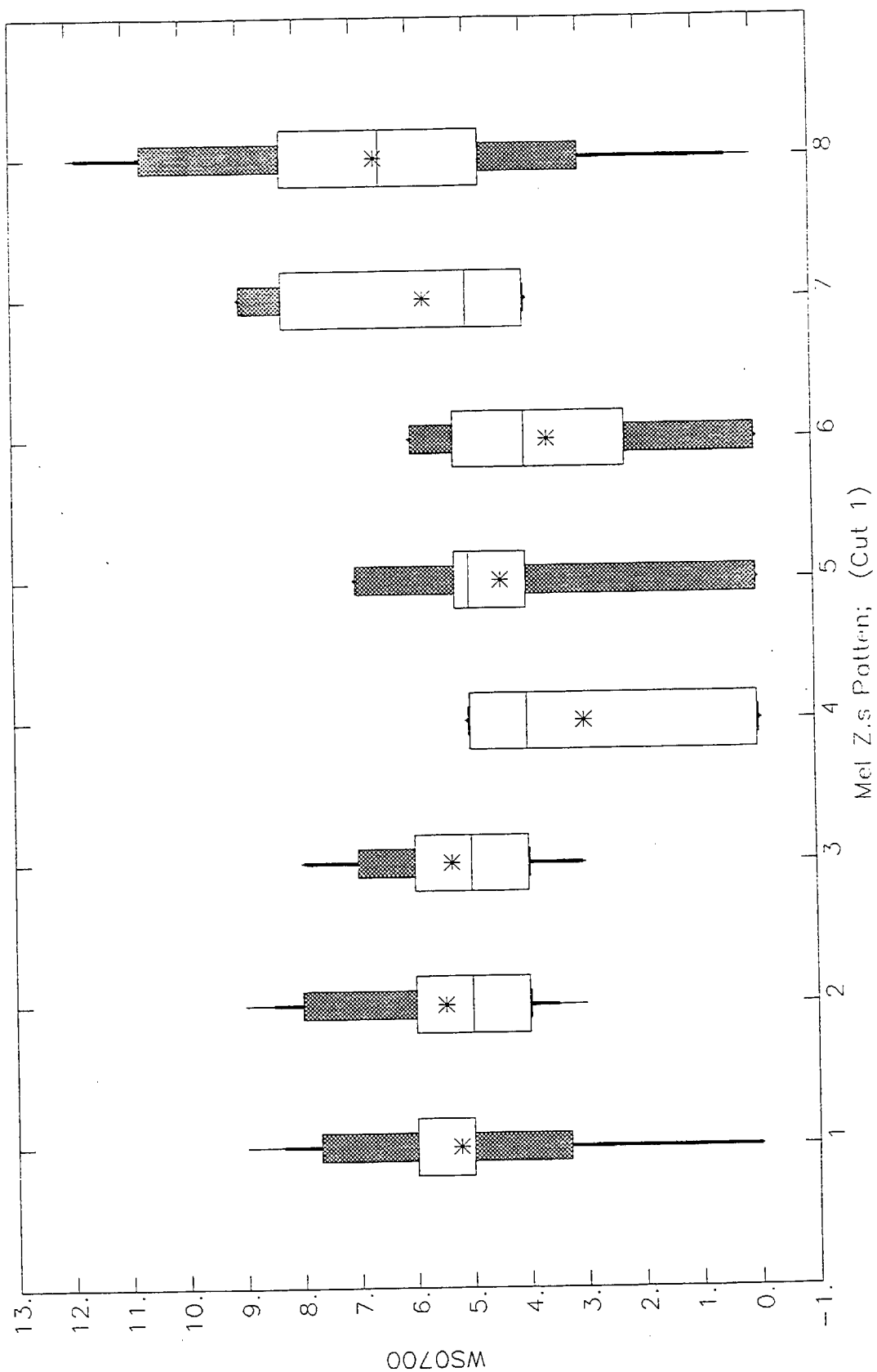


FIGURE 3-6. Boxplot of variable WS0700 for each source history category. (Key to boxplots appears in Figure 3-3.)

Box plot for variable WS1000, *==mean
 order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

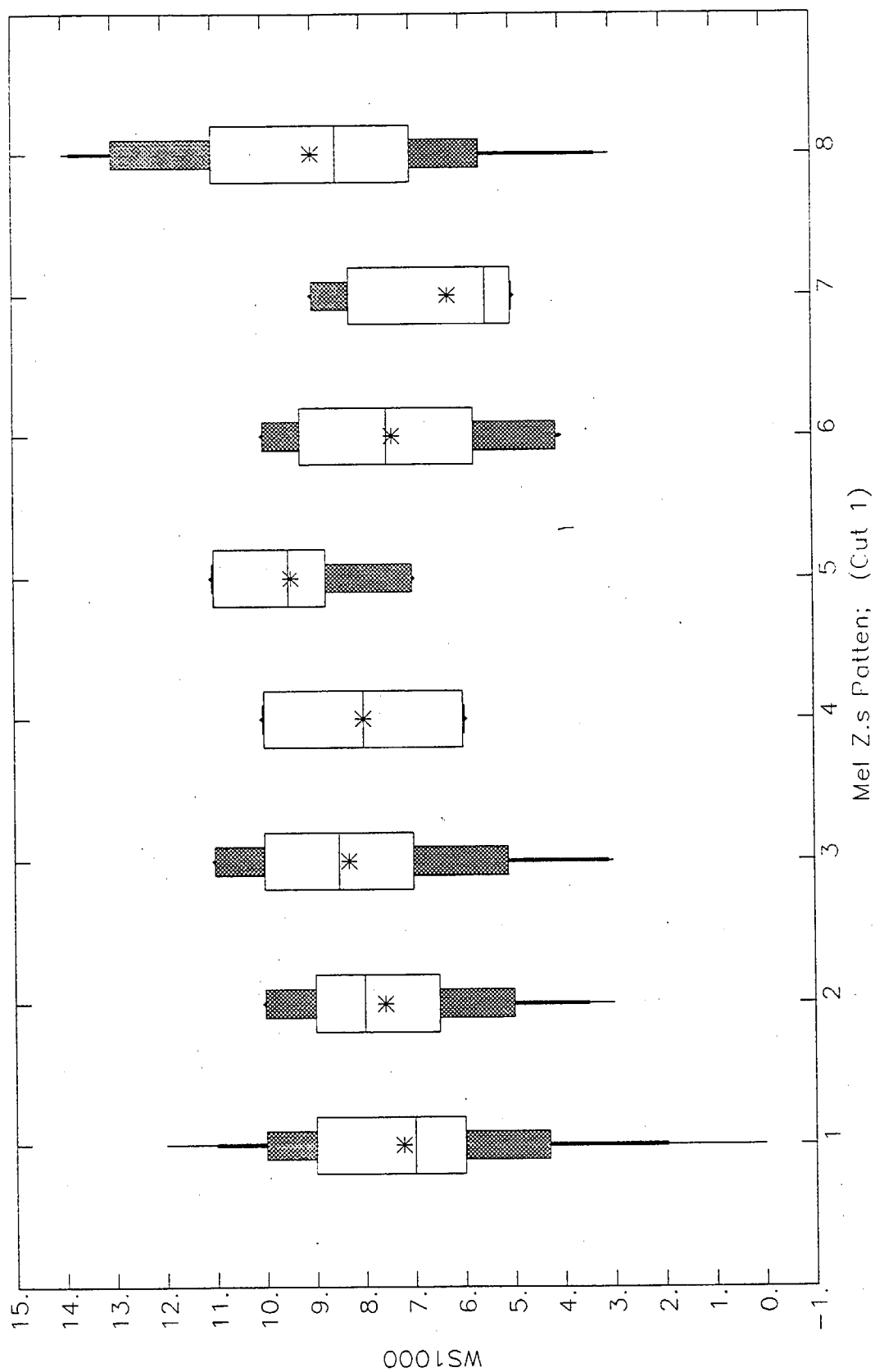


FIGURE 3-7. Boxplot of variable WS1000 for each source history category. (Key to boxplots appears in Figure 3-3.)

Box plot for variable SKY0700, *=mean
order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

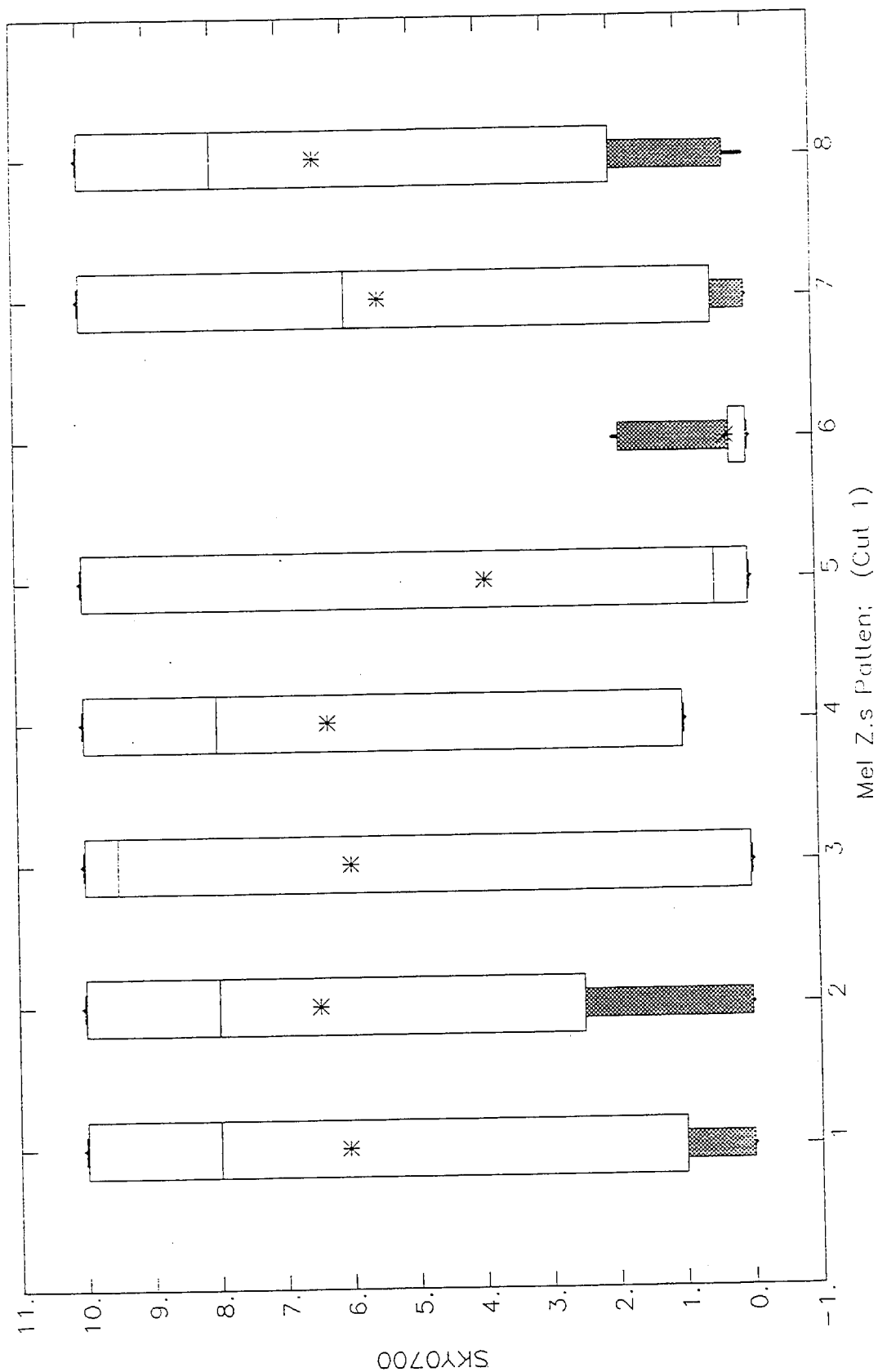


FIGURE 3-8. Boxplot of variable SKY0700 for each source history category. (Key to boxplots appears in Figure 3-3.)

Box plot for variable CELL0700, *=mean
 order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

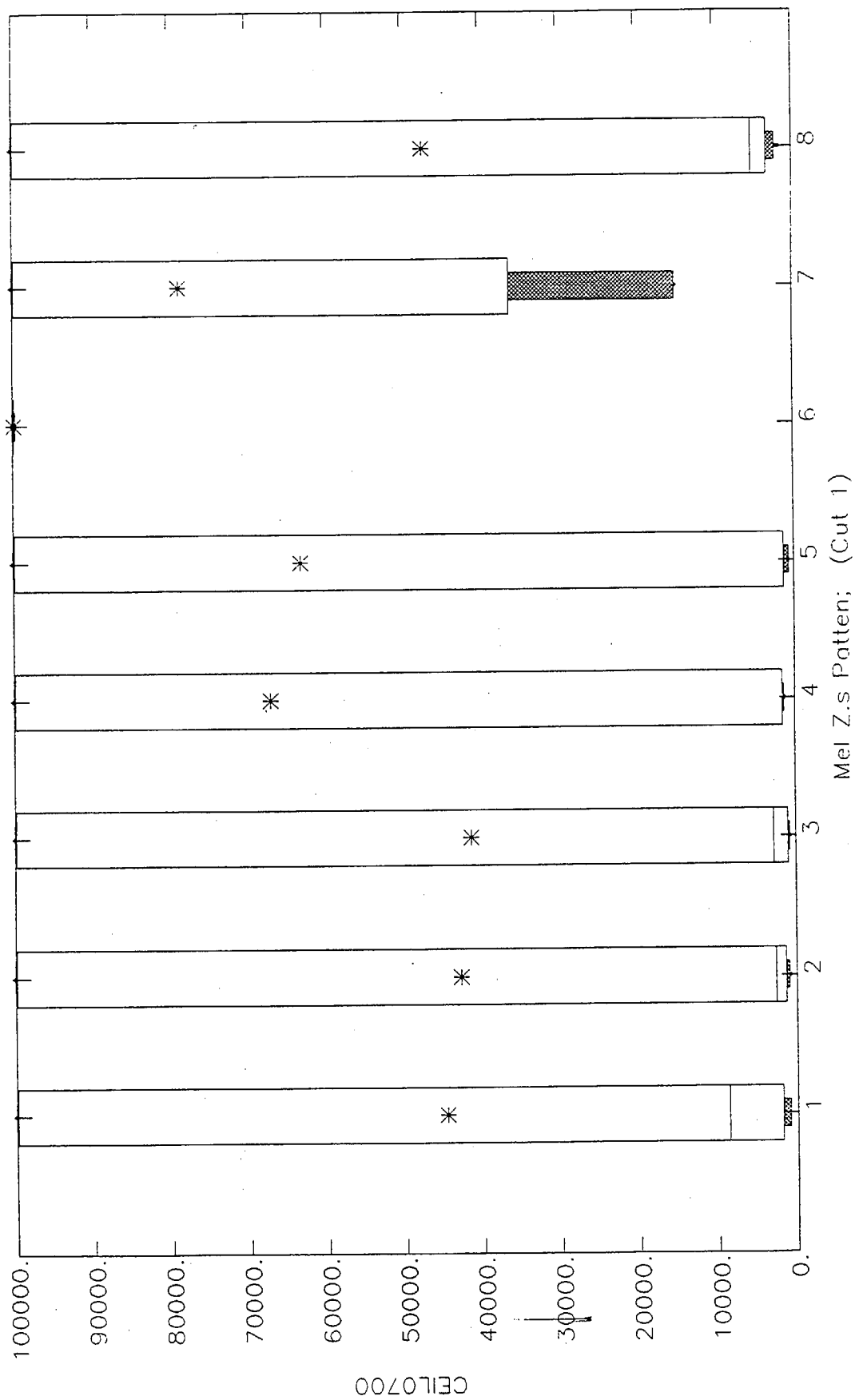


FIGURE 3-9. Boxplot of variable CELL0700 for each source history category. (Key to boxplots appears in Figure 3-3.)

Box plot for variable SKY1000, * = mean
 order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

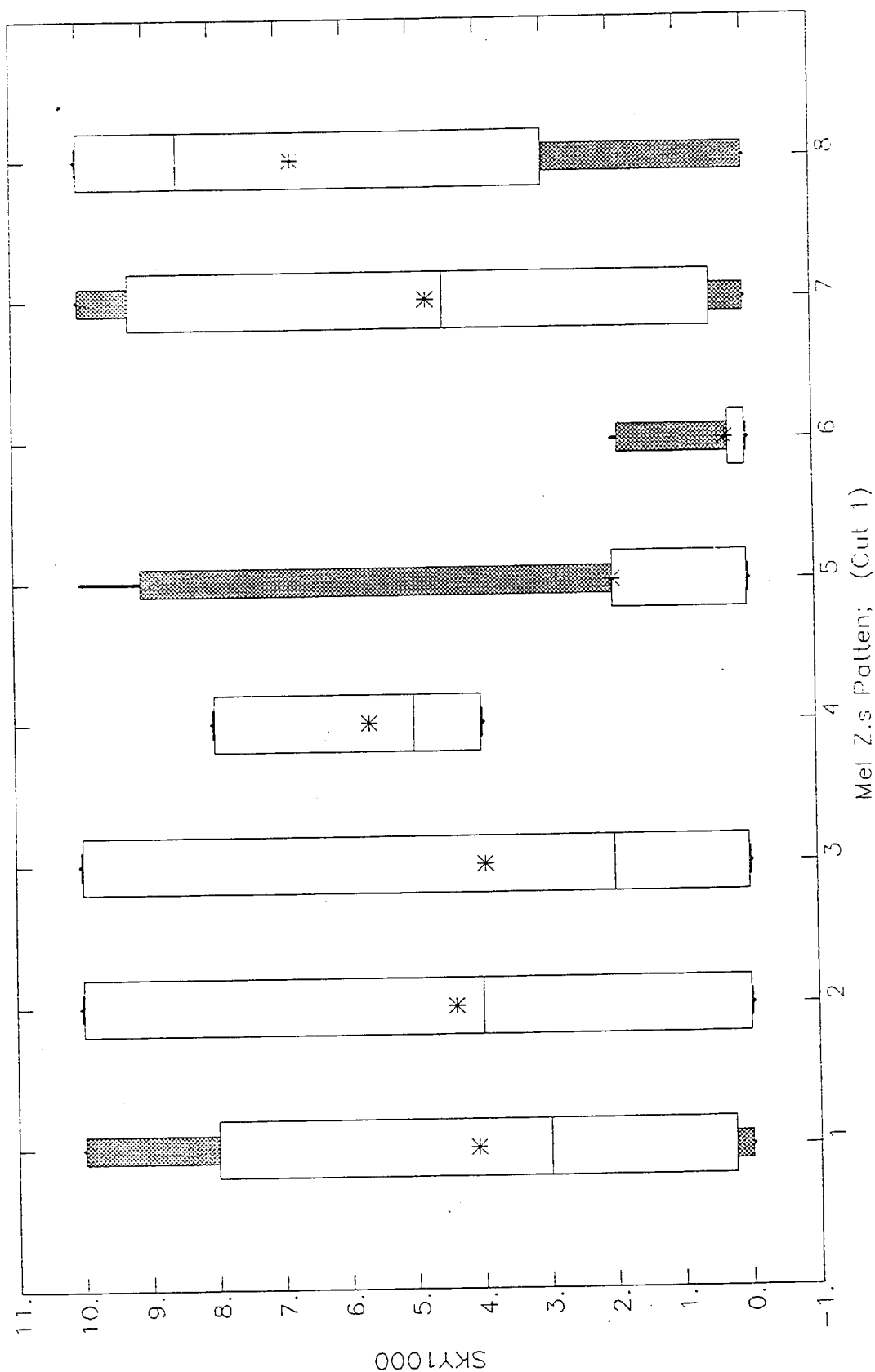


FIGURE 3-10. Boxplot of variable SKY1000 for each source history category. (Key to boxplots appears in Figure 3-3.)

Box plot for variable CEIL1000, +=mean
 order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

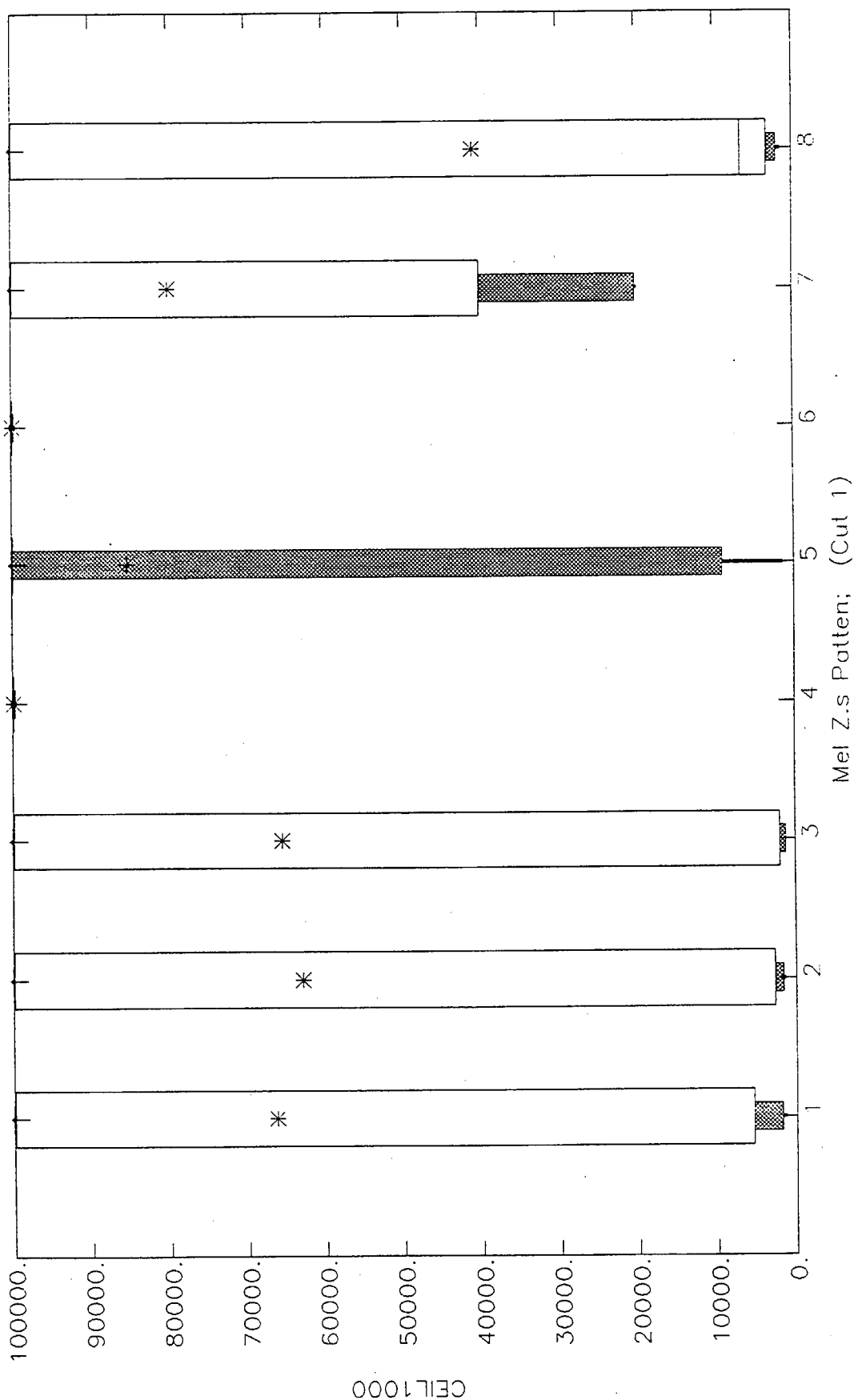


FIGURE 3-11. Boxplot of variable CEIL1000 for each source history category. (Key to boxplots appears in Figure 3-3.)

Box plot for variable LATHPRO7, * = mean
order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

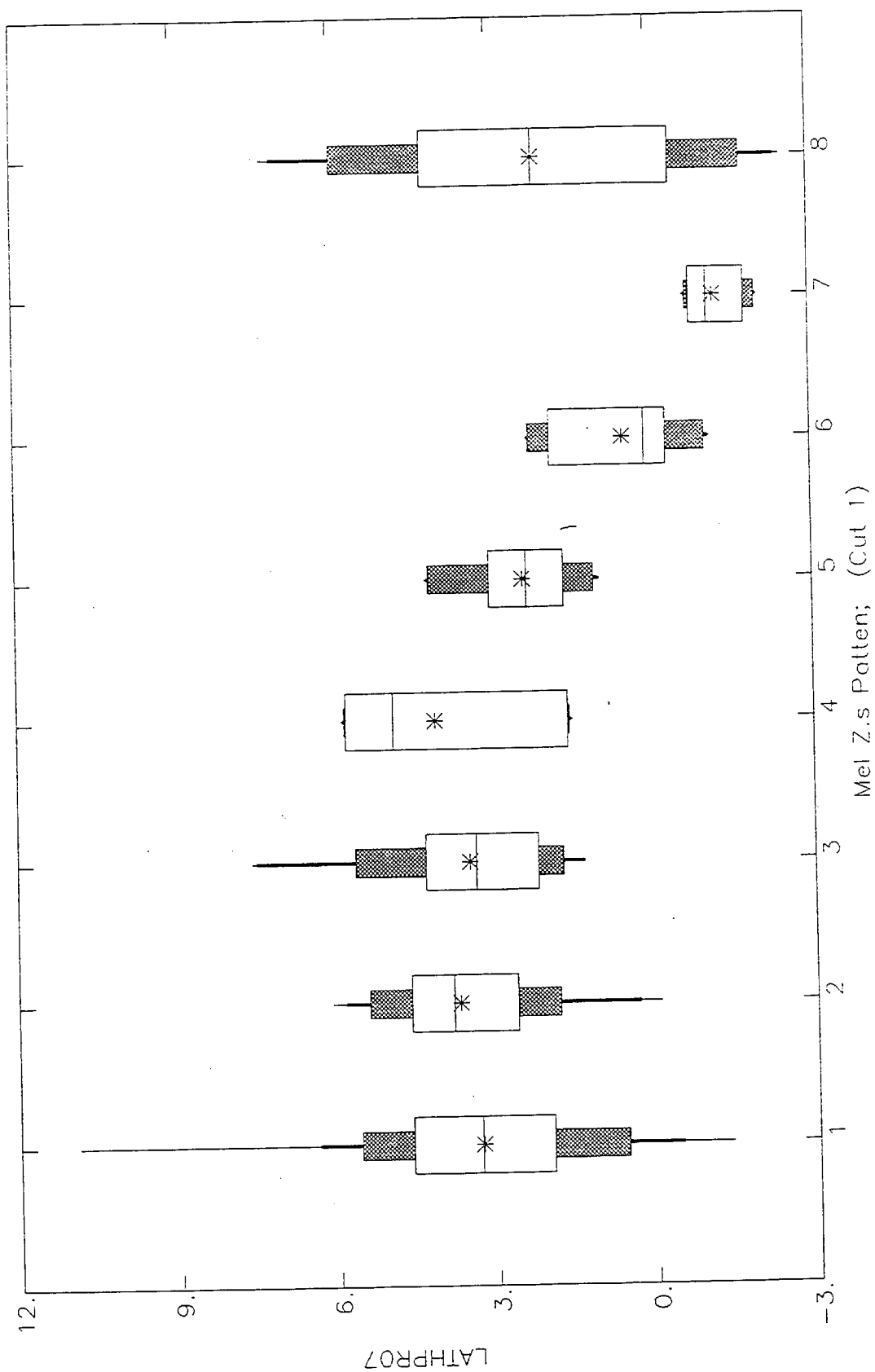


FIGURE 3-12. Boxplot of variable LATHPRO7 for each source history category. (Key to boxplots appears in Figure 3-3.)

Box plot for variable LATHPR16, *=mean
order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

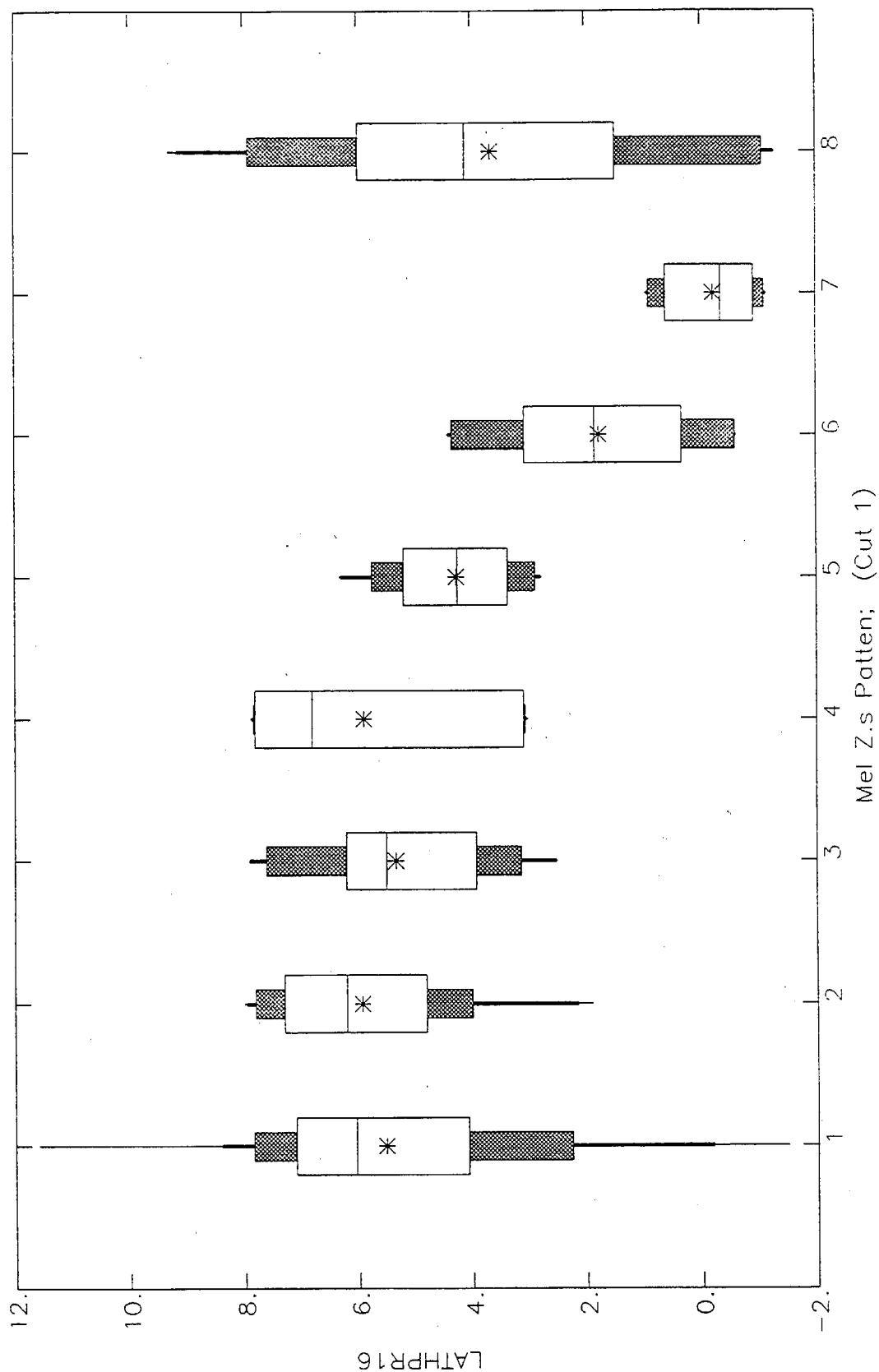


FIGURE 3-13. Boxplot of variable LATHPR16 for each source history category. (Key to boxplots appears in Figure 3-3.)

Box plot for variable PGLAXWJF, * = mean
 order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

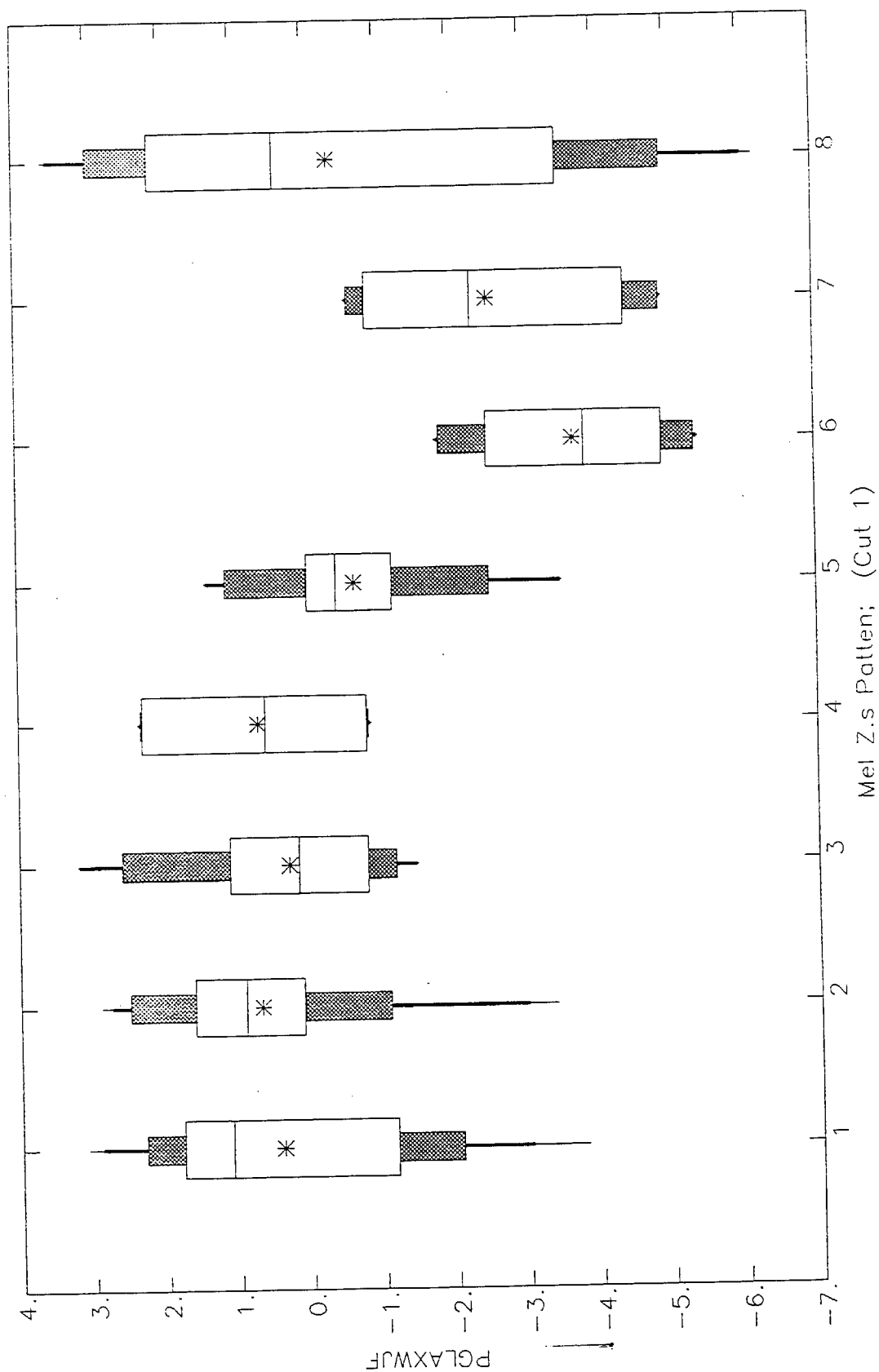


FIGURE 3-14. Boxplot of variable PGLAXWJF for each source history category. (Key to boxplots appears in Figure 3-3.)

Box plot for variable AWSLGB, * = mean
 order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

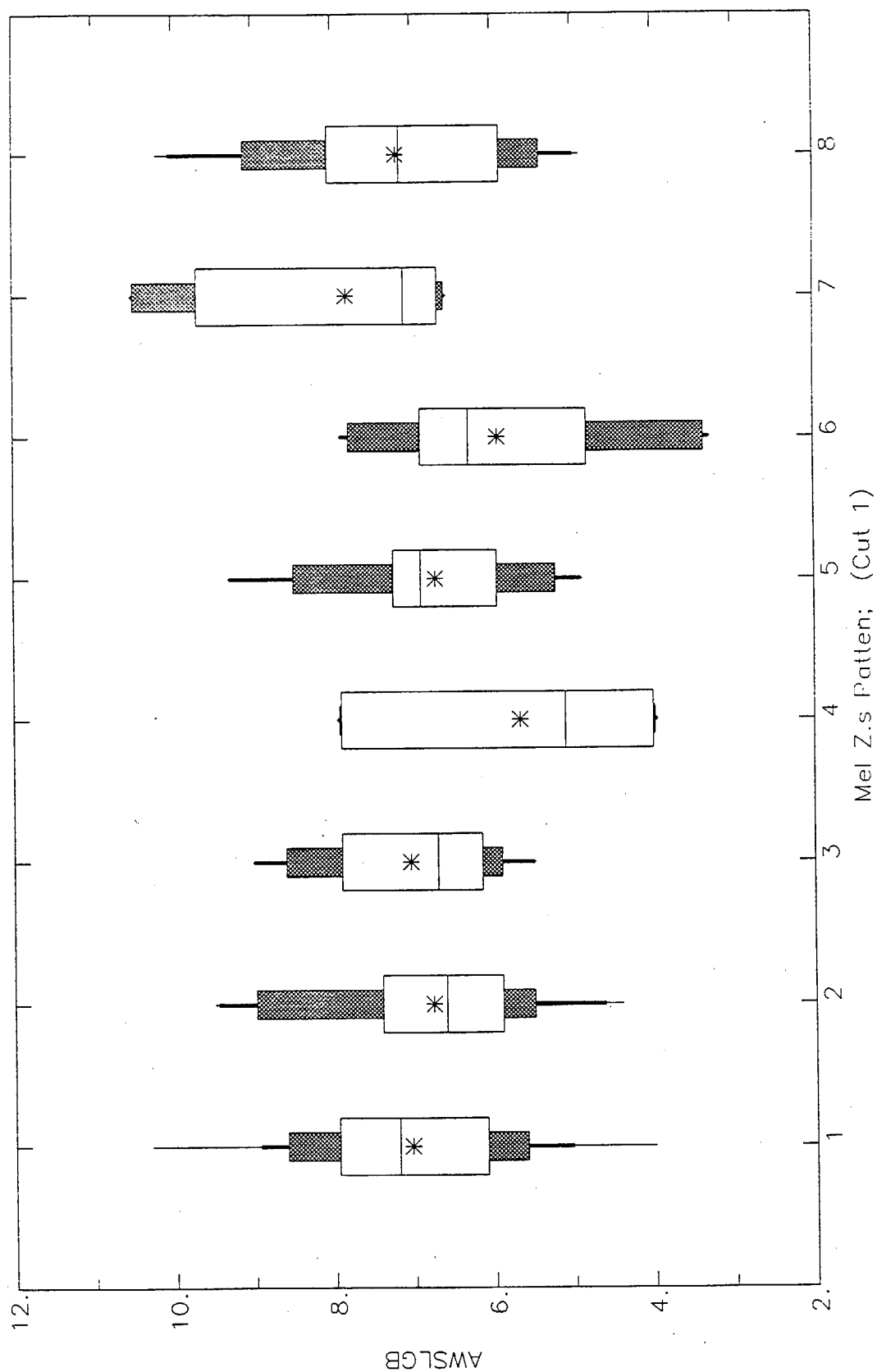


FIGURE 3-15. Boxplot of variable AWSLGB for each source history category. (Key to boxplots appears in Figure 3-3.)

Box plot for variable RWSLGB, *==mean
 order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

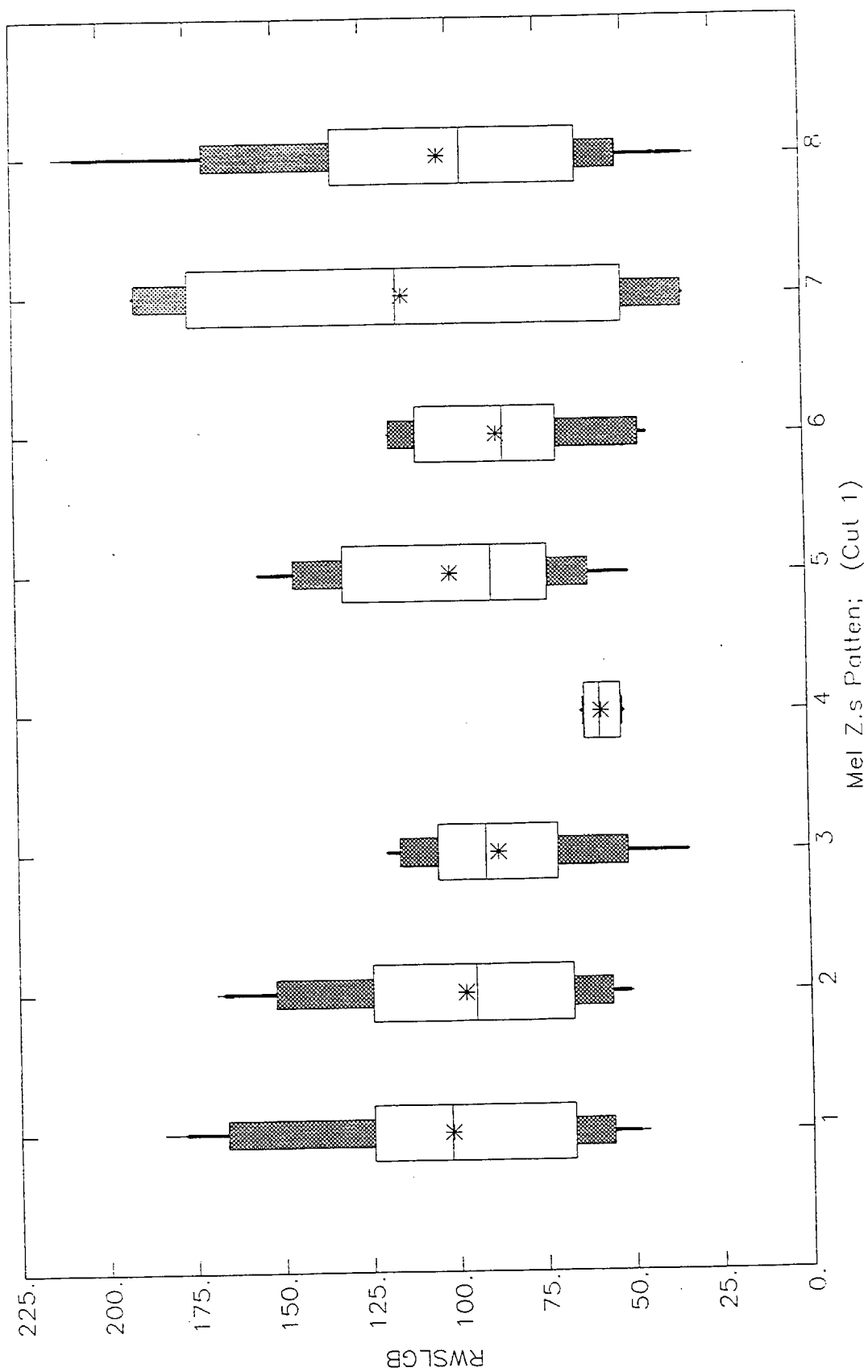


FIGURE 3-16. Boxplot of variable RWSLGB for each source history category. (Key to boxplots appears in Figure 3-3.)

Box plot for variable U0700, *=mean
order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

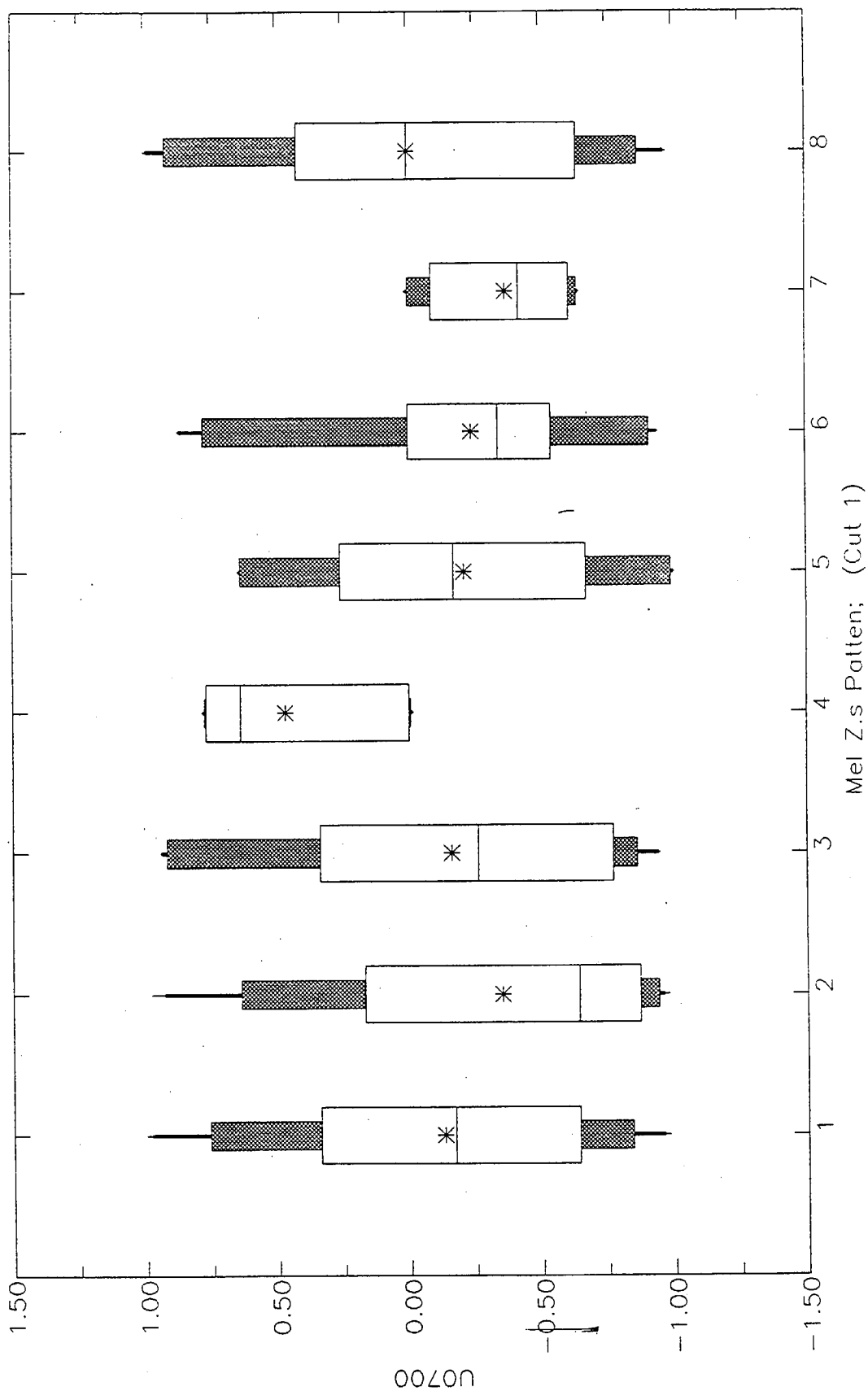


FIGURE 3-17. Boxplot of variable U0700 for each source history category. (Key to boxplots appears in Figure 3-3.)

Box plot for variable V0700, *=mean
 order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

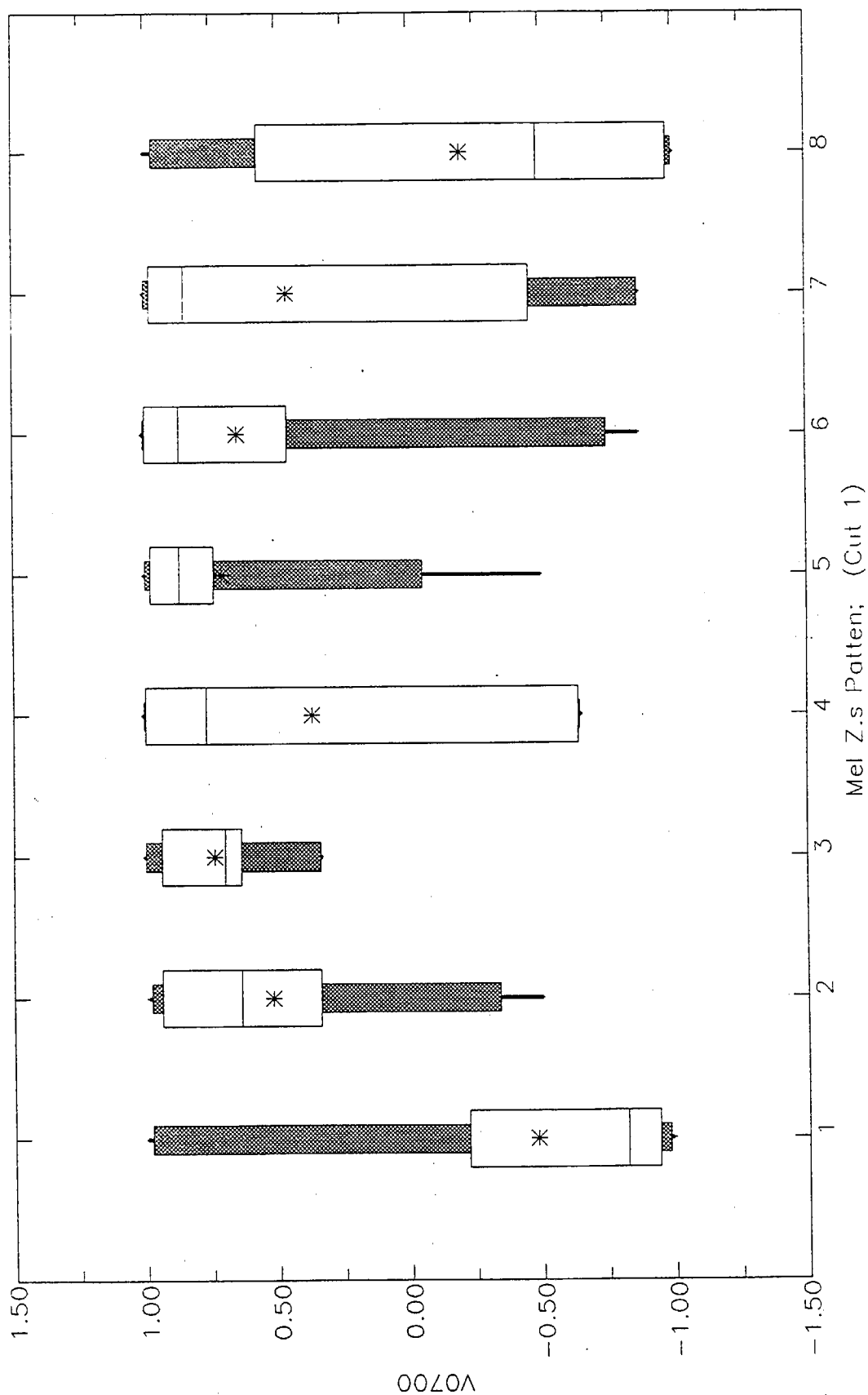


FIGURE 3-18. Boxplot of variable V0700 for each source history category. (Key to boxplots appears in Figure 3-3.)

Box plot for variable U1000, *=mean
order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

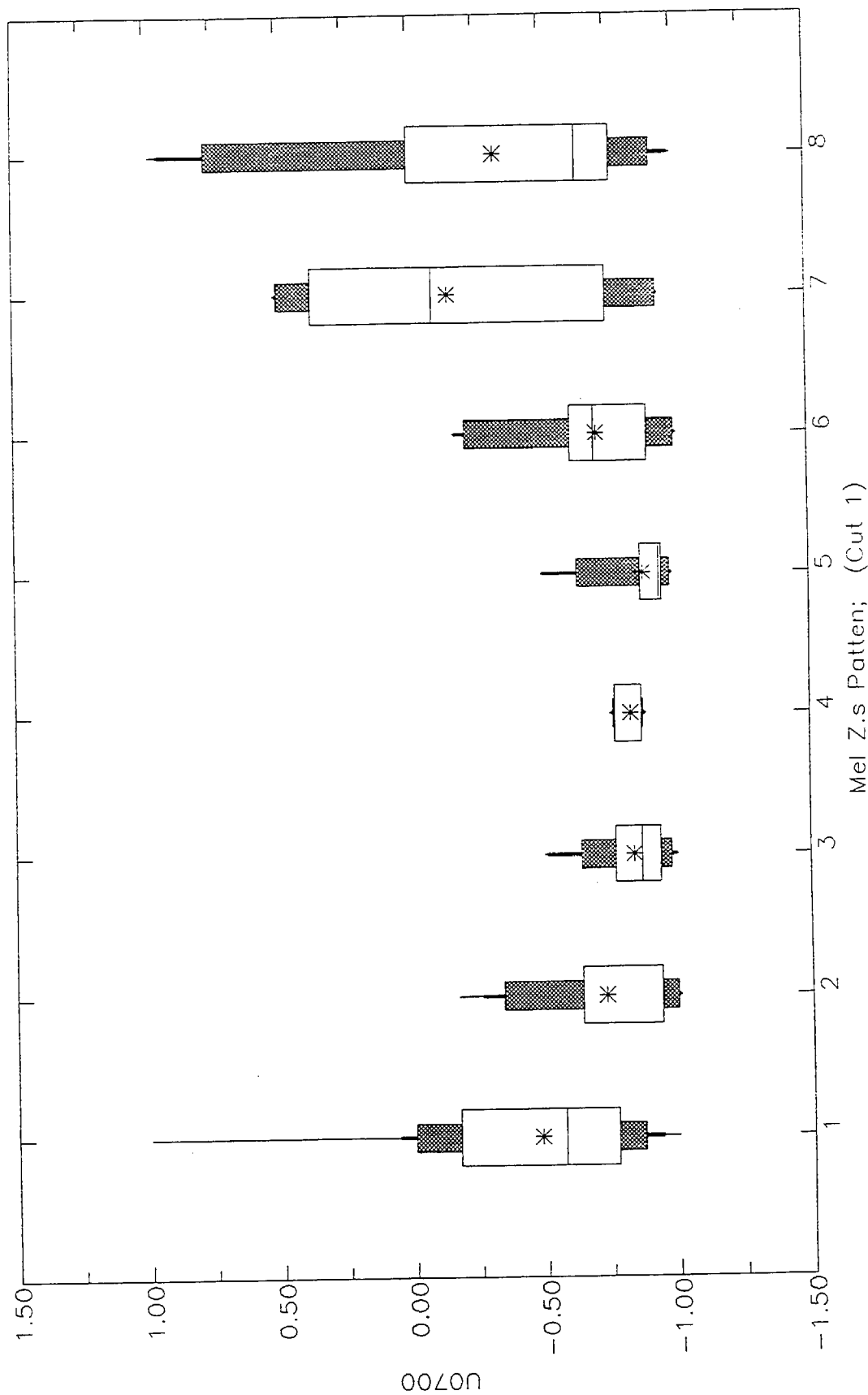


FIGURE 3-19. Boxplot of variable U1000 for each source history category. (Key to boxplots appears in Figure 3-3.)

Box plot for variable V1000, *=mean
 order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

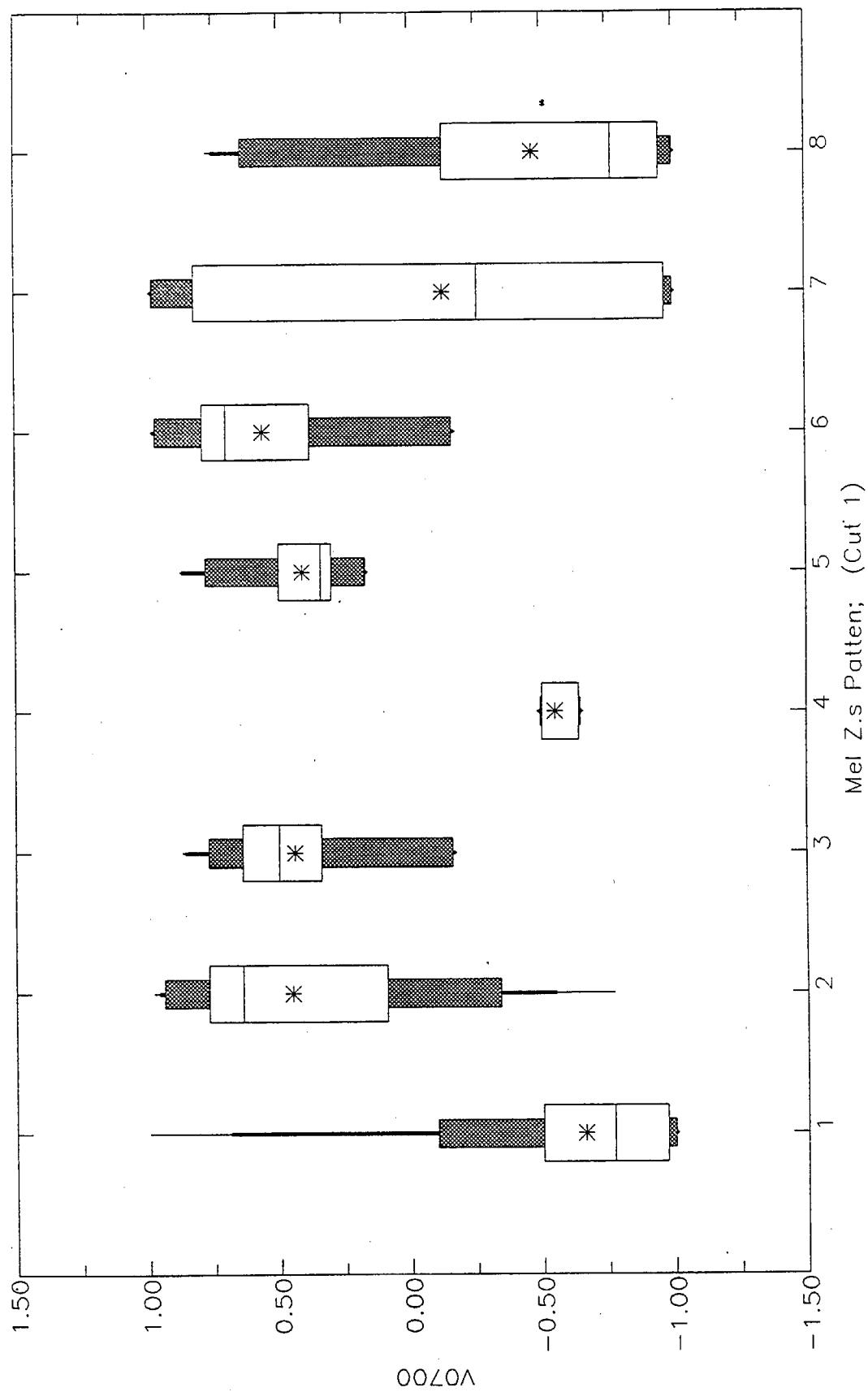


FIGURE 3-20. Boxplot of variable V1000 for each source history category. (Key to boxplots appears in Figure 3-3.)

Box plot for variable ULGB, * = mean
 order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

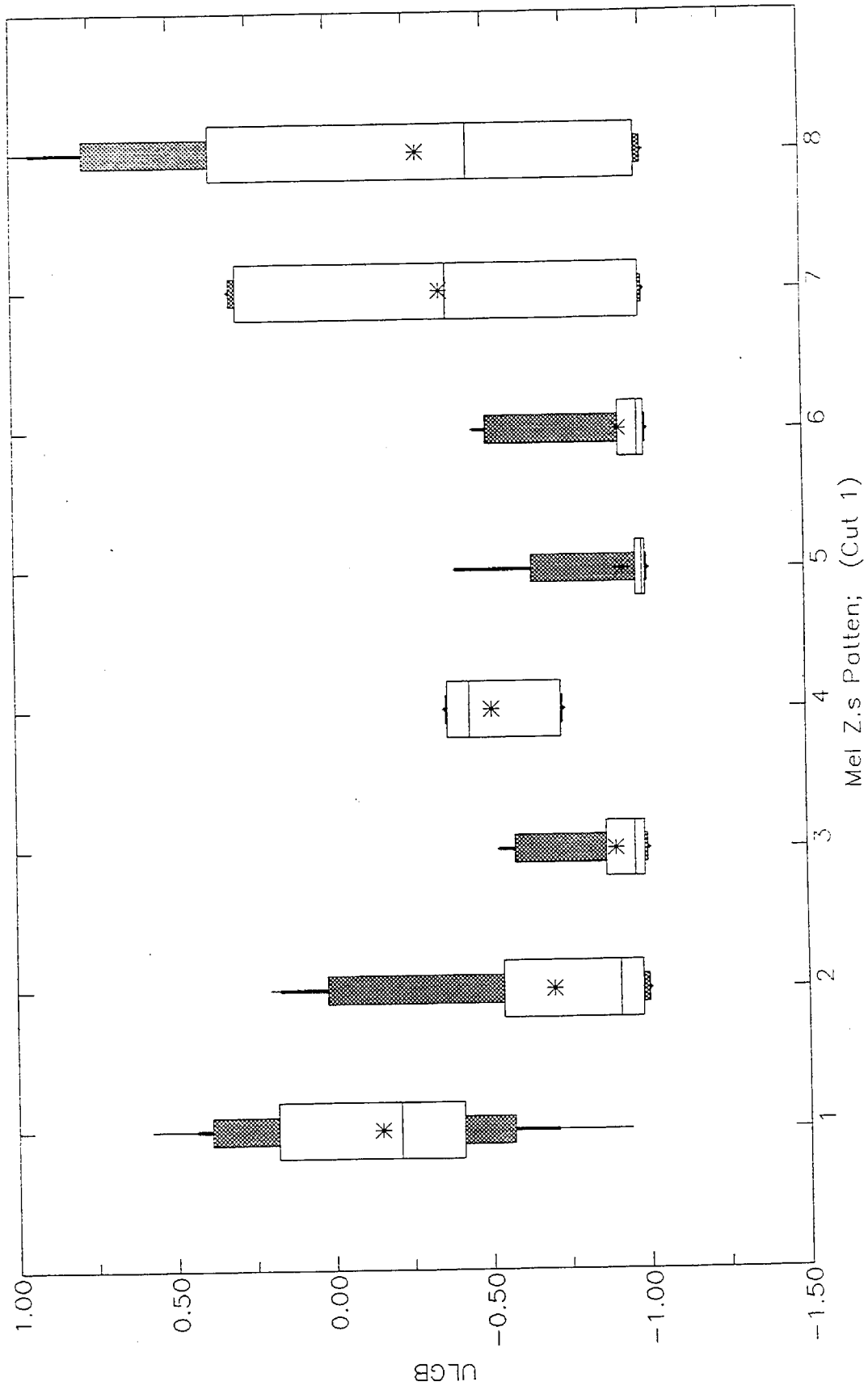


FIGURE 3-21. Boxplot of variable ULGB for each source history category. (Key to boxplots appears in Figure 3-3.)

Box plot for variable VLGB, * = mean
 order = min, 5%, 10%, 25%, 50%, 75%, 90%, 95%, max.

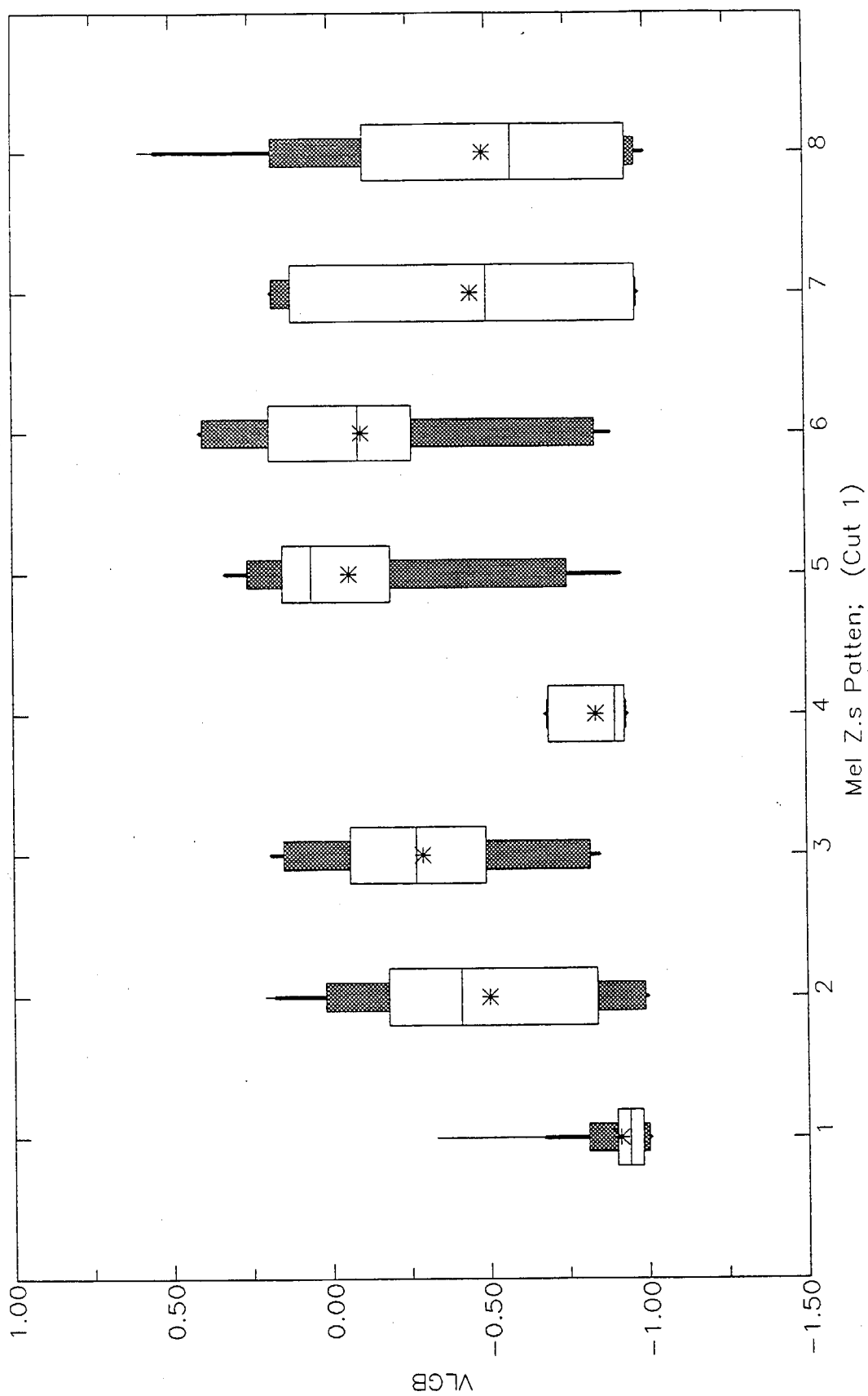


FIGURE 3-22. Boxplot of variable VLGB for each source history category. (Key to boxplots appears in Figure 3-3.)

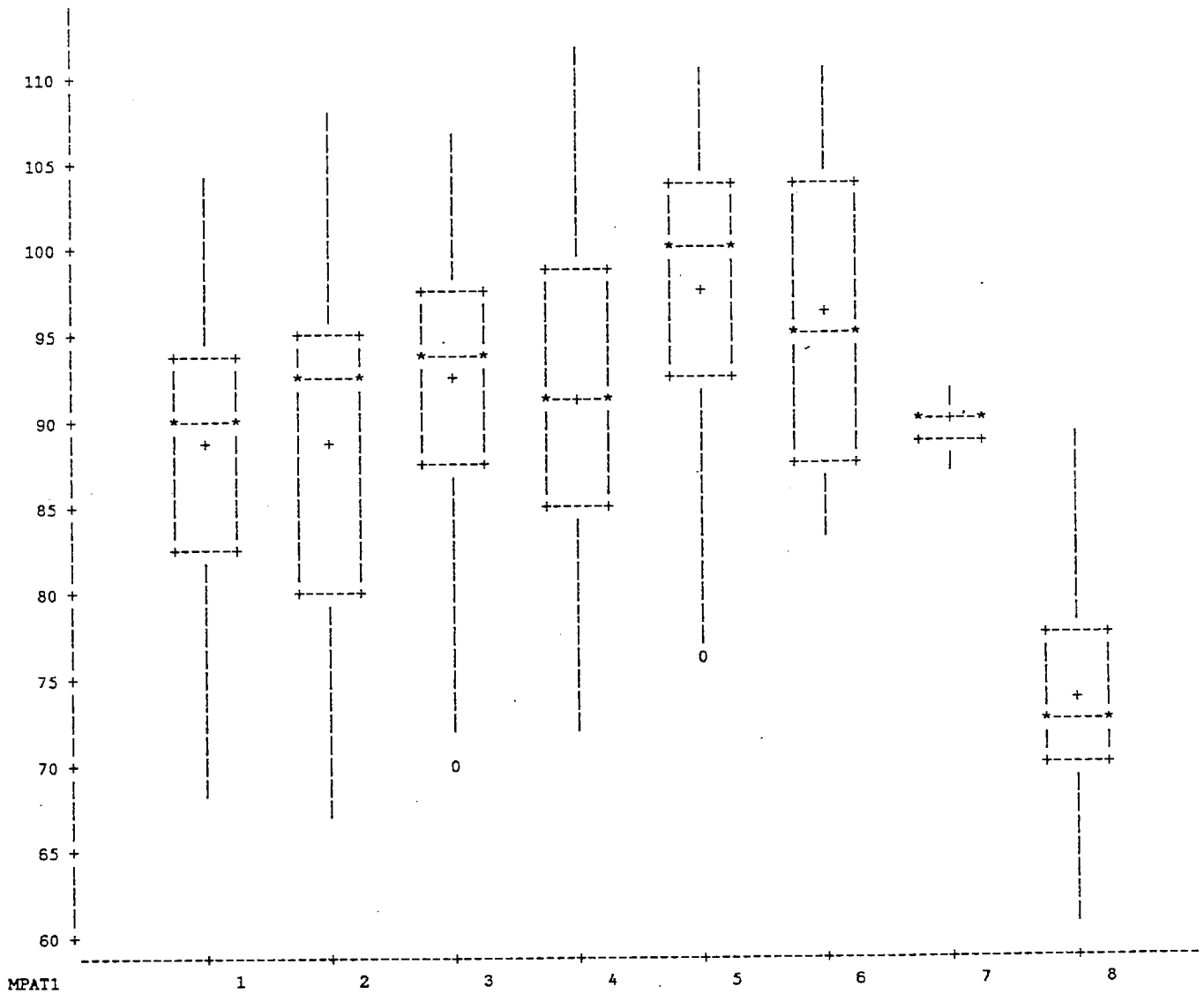


FIGURE 3-23. Box plots of variable TMAXSBD for each source history category. (Key to box plots is presented in Figure 3-30.)

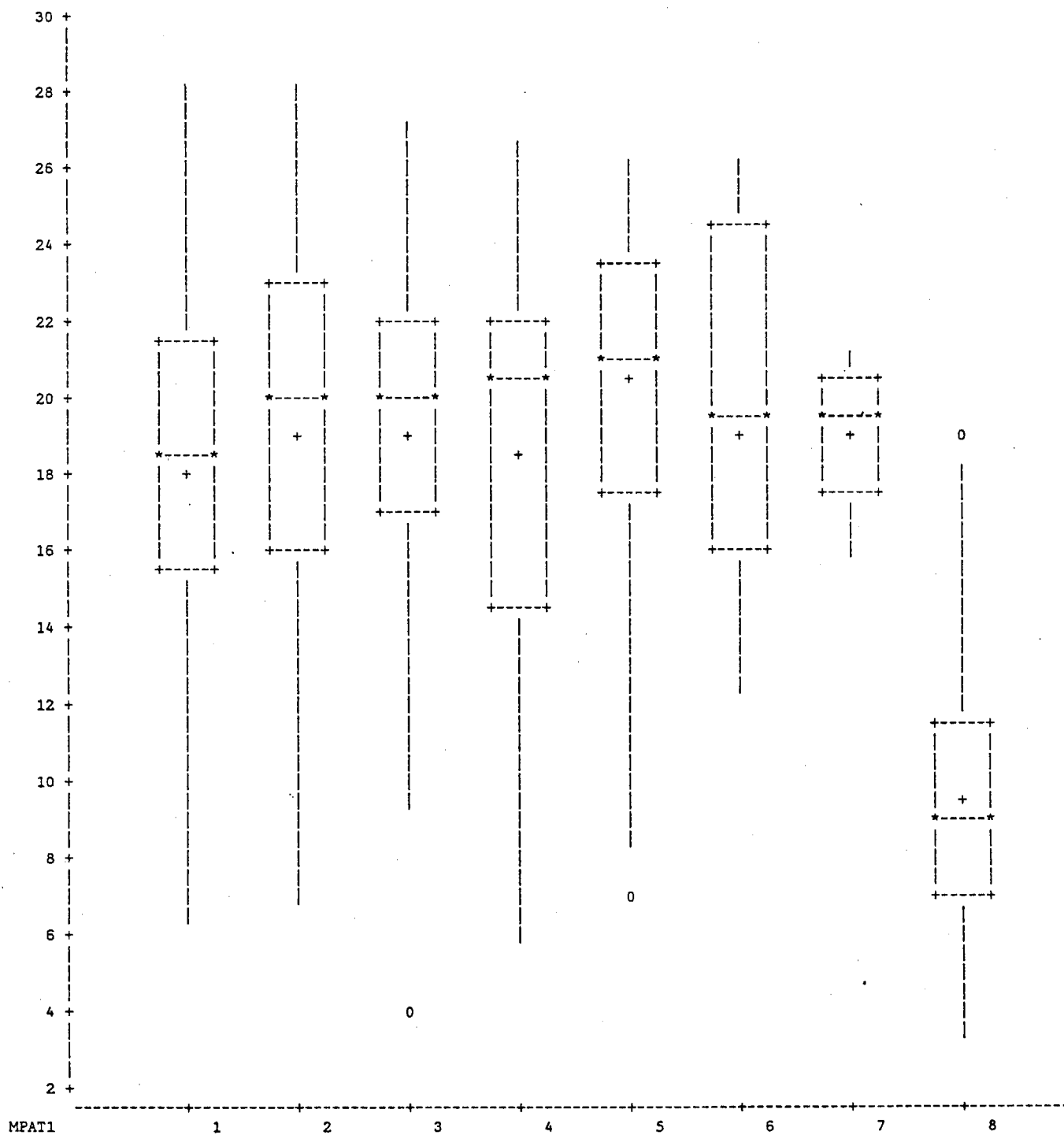


FIGURE 3-24. Box plots of variable T850 for each source history category. (Key to box plots is presented in Figure 3-30.)

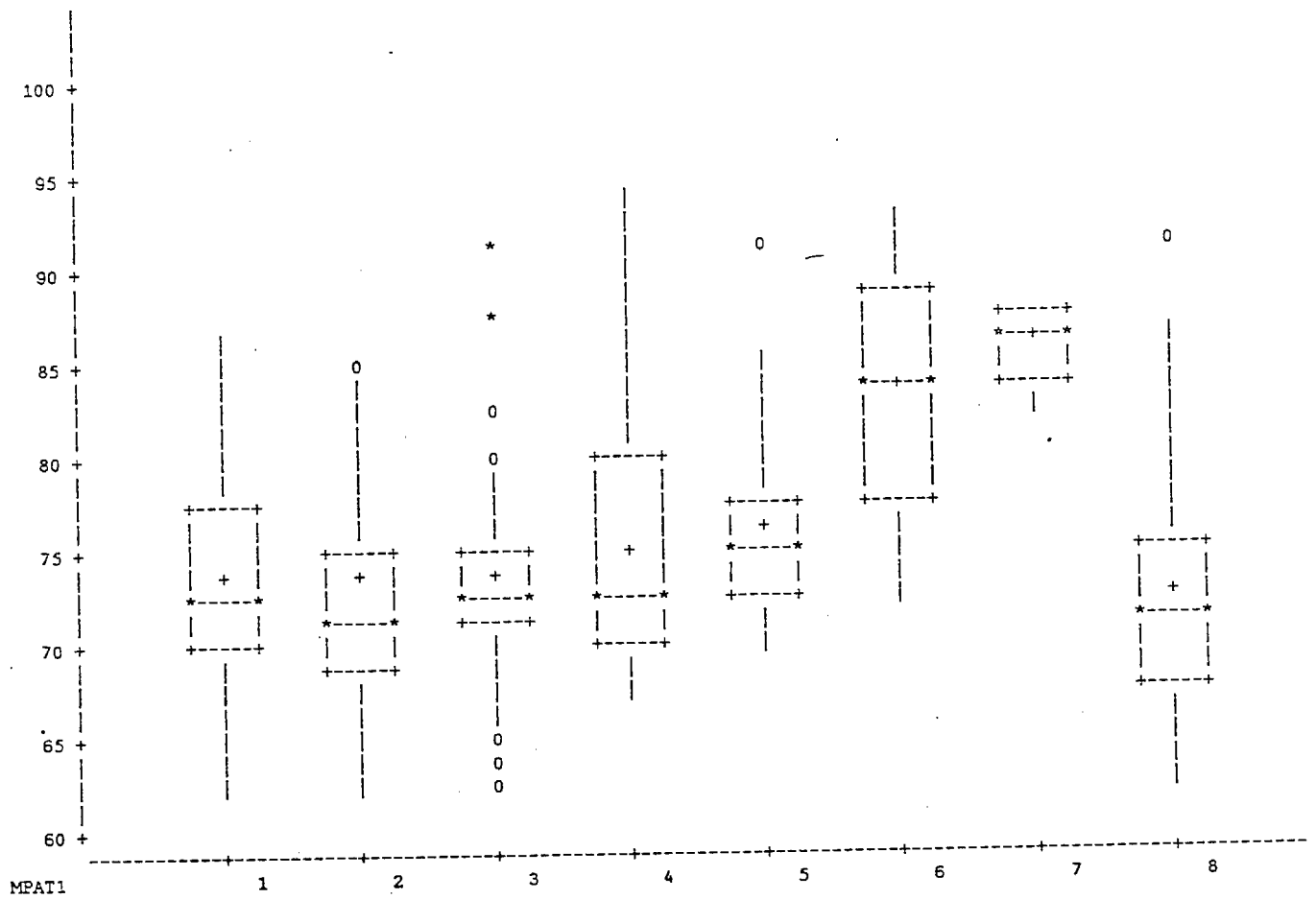


FIGURE 3-25. Box plots of variable TMAXLAX for each source history category. (Key to box plots is presented in Figure 3-30.)

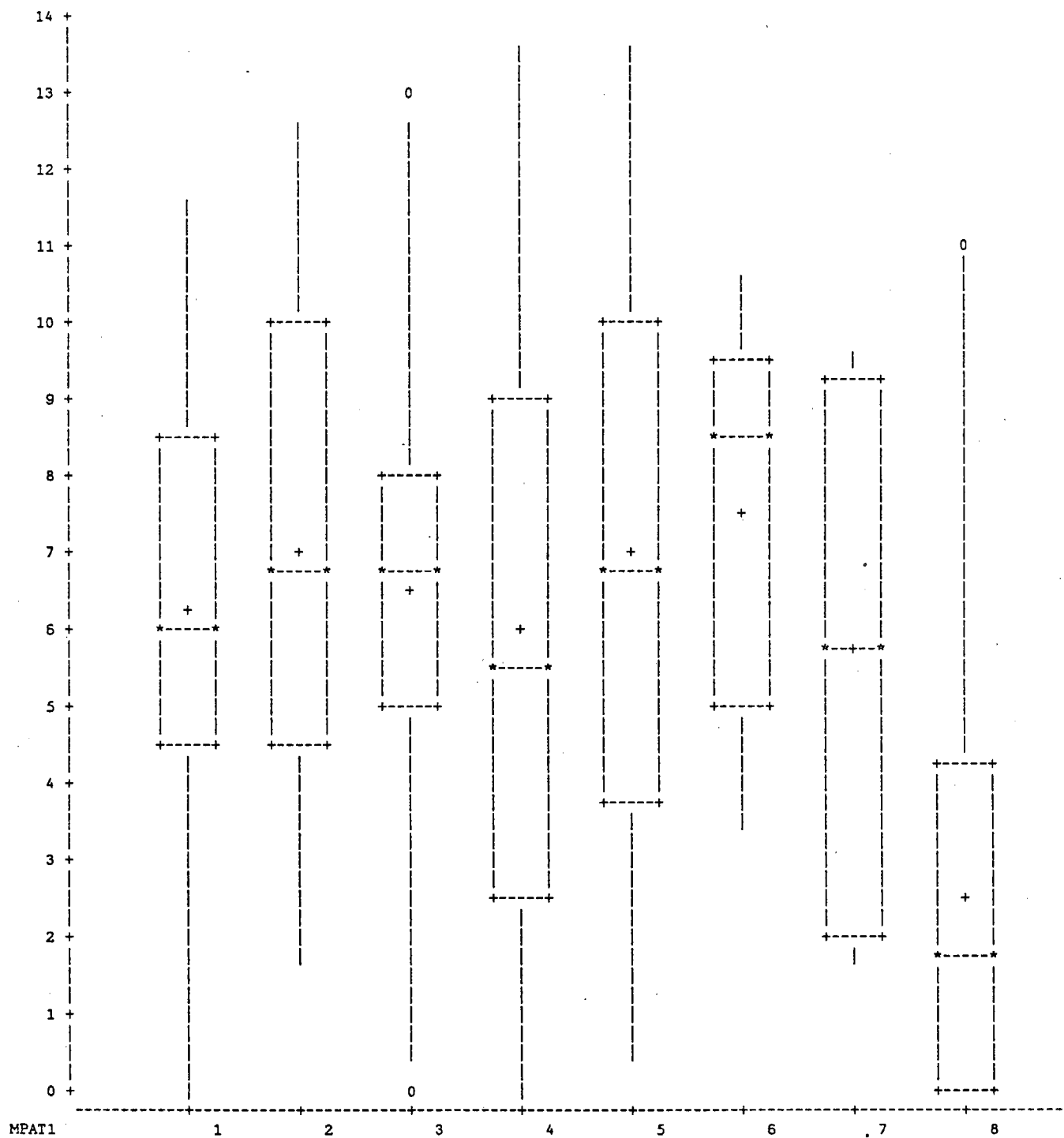


FIGURE 3-26. Box plots of variable DELTAT for each source history category. (Key to box plots is presented in Figure 3-30.)

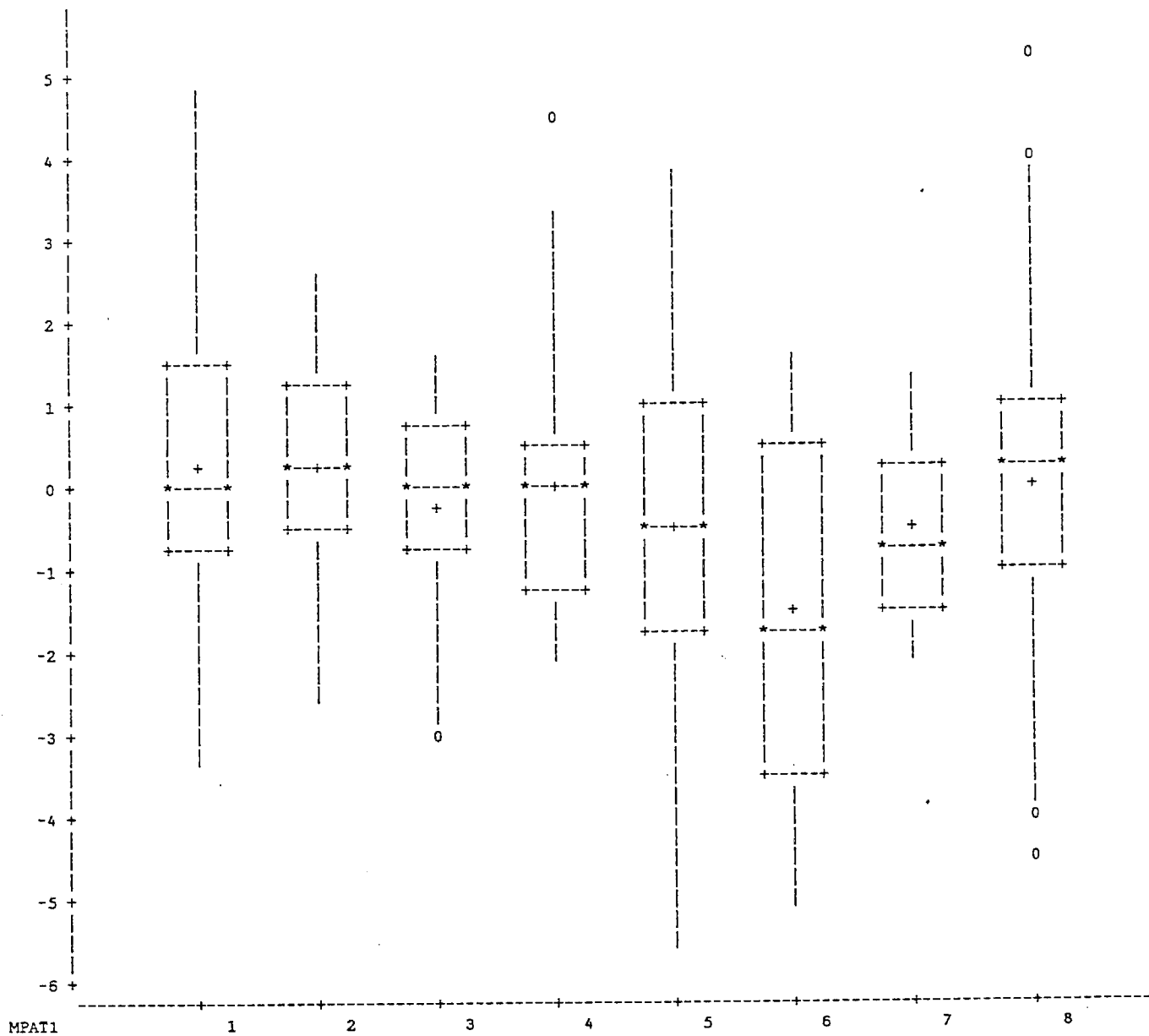


FIGURE 3-27. Box plots of variable DELTAPG for each source history category. (Key to box plots is presented in Figure 3-30.)

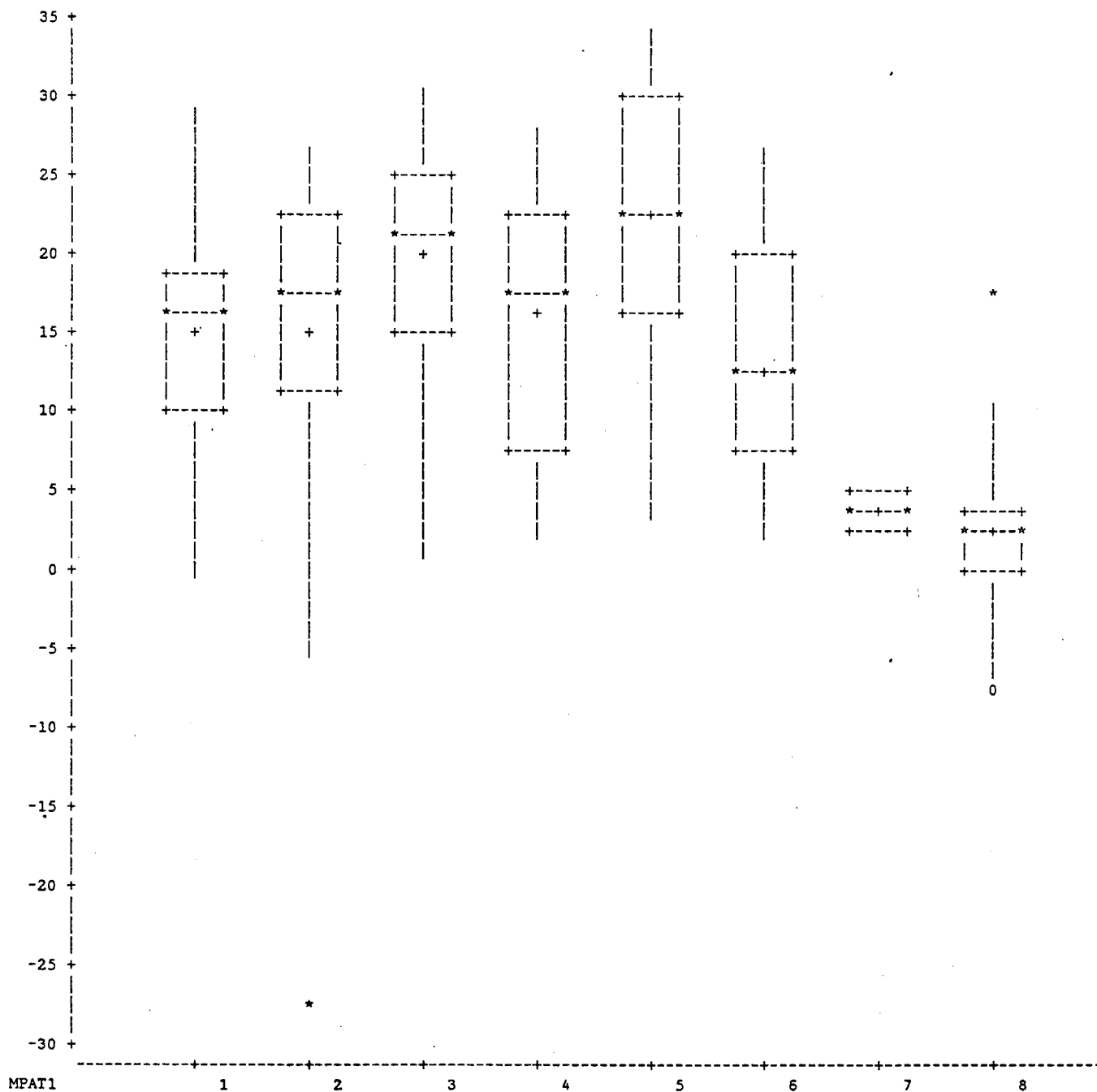


FIGURE 3-28. Box plots of variable MAXDIFF1 for each source history category. (Key to box plots is presented in Figure 3-30.)

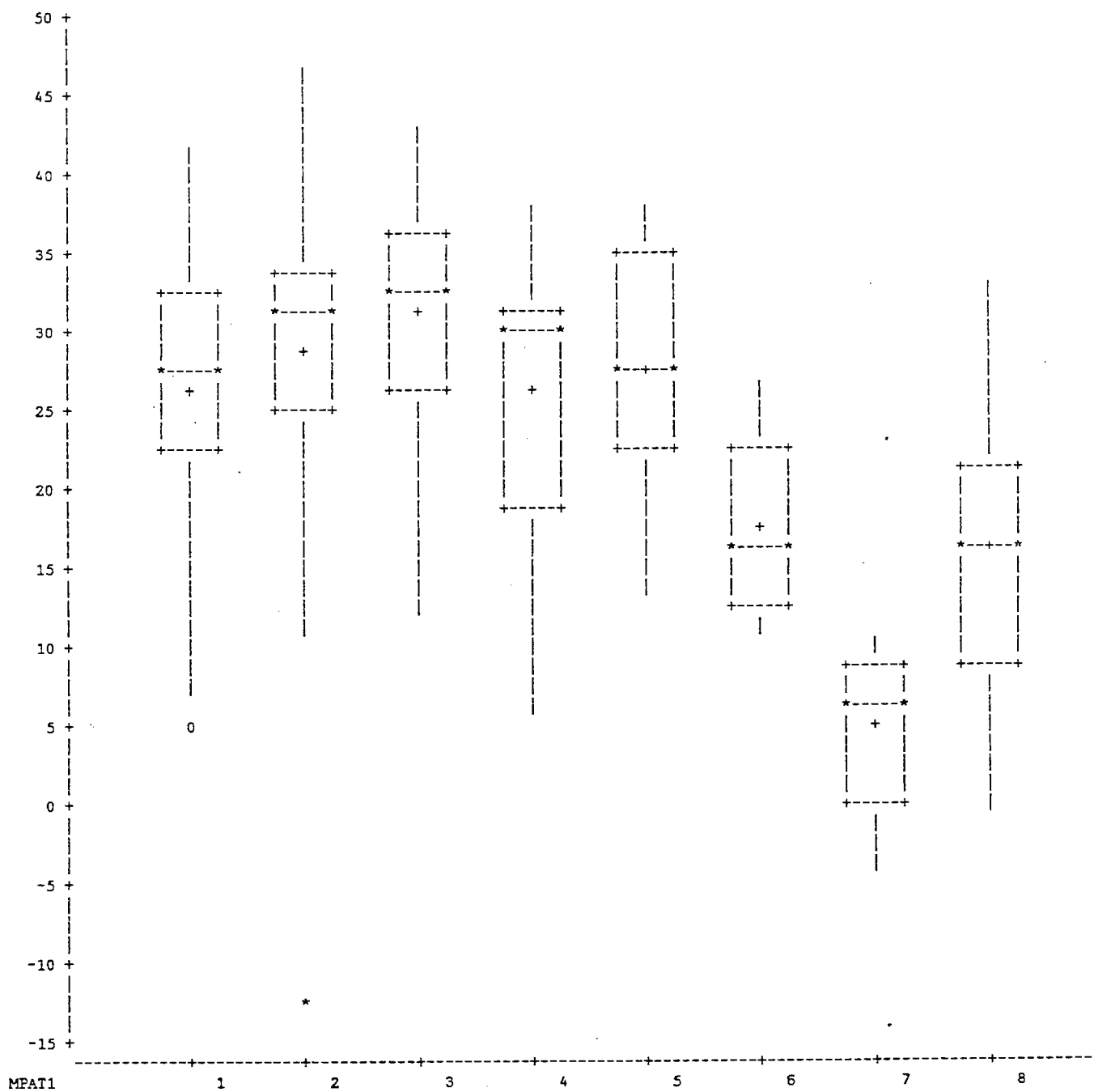


FIGURE 3-29. Box plots of variable MAXDIFF2 for each source history category. (Key to box plots is presented in Figure 3-30.)

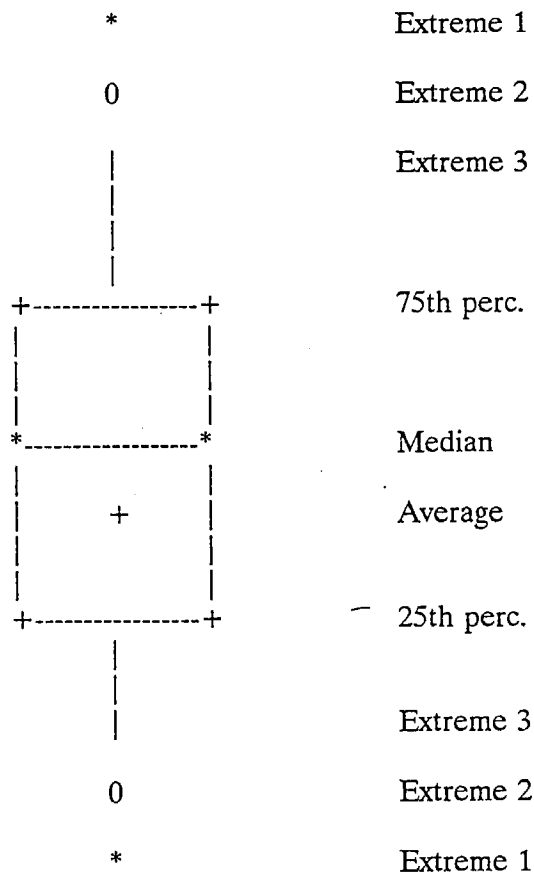


FIGURE 3-30. Key for boxplots in Figures 3-23 to 3-29; Extreme 1 = values more than 3 times interquartile range* above or below median. Extreme 2 = values between 1.5 and 3 times interquartile range above or below median. Extreme 3 = largest (smallest) value that is not more than 1.5 times interquartile range above (below) median.

* Interquartile range is the interval between the 75th and 25th percentiles.

Inversion Base Heights by Category

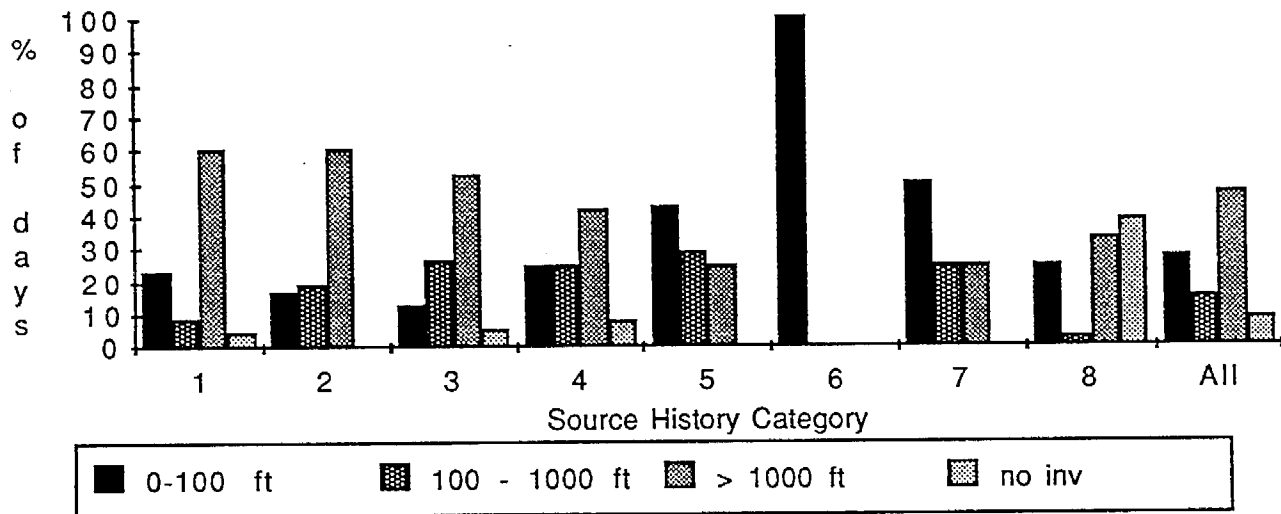


FIGURE 3-31. Distributions of inversion base height elevations from 1300 UTC LMU/UCLA sounding (BASEHT) for days in each source history category.

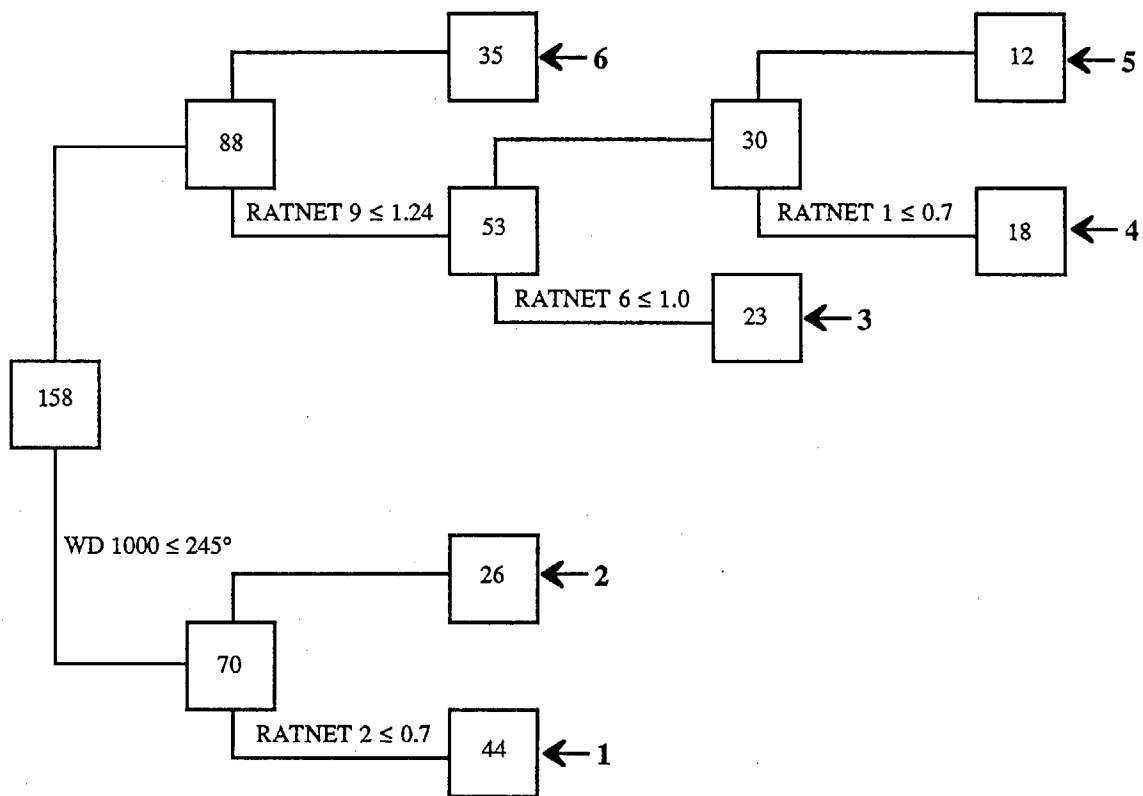


FIGURE 3-32. Classification tree illustrating relationship between ozone patterns and meteorological and aerometric variables. Numbers in boxes indicate number of days at each node; variable names are defined in the text. Bold numbers to right identify terminal nodes.

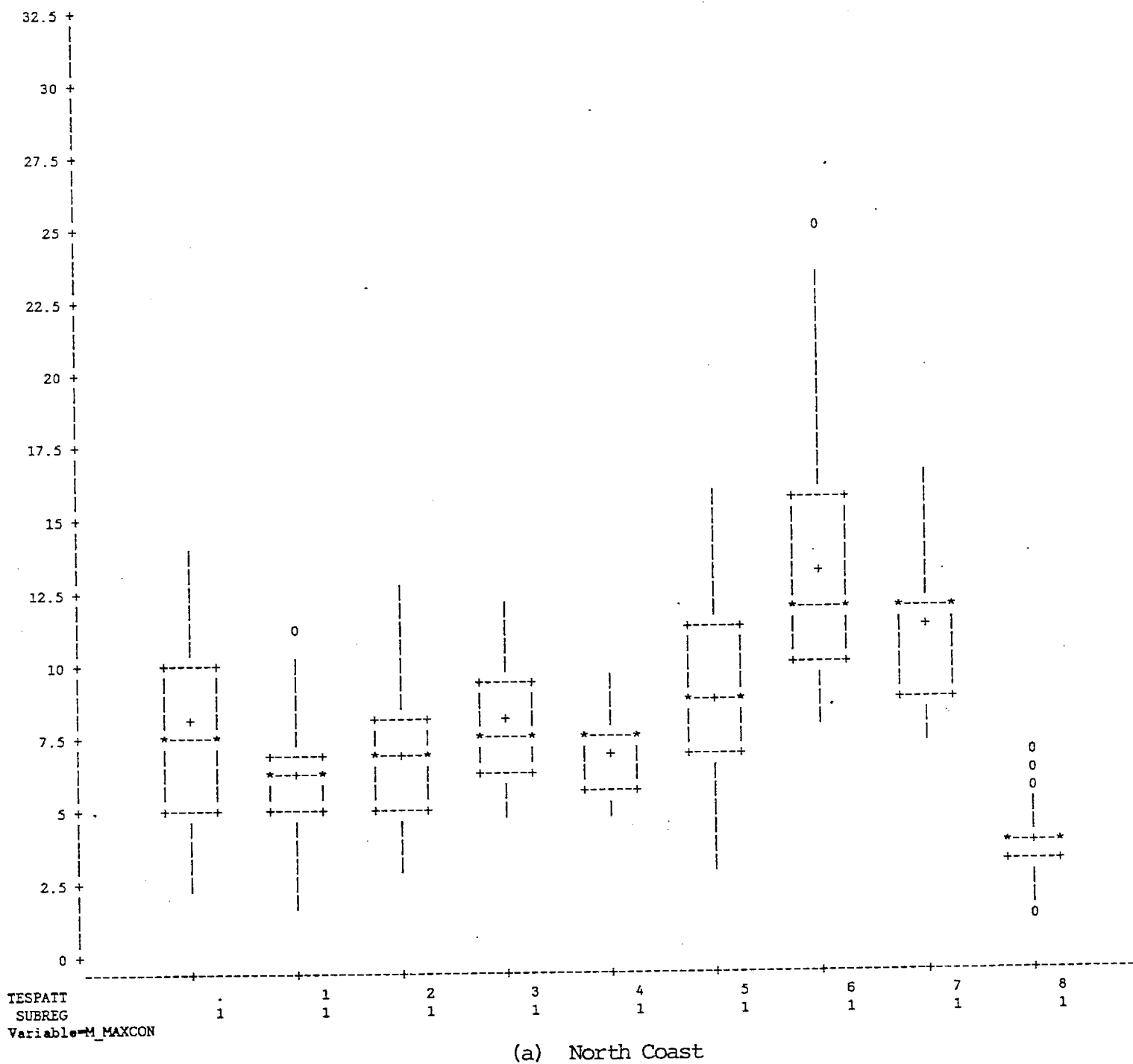


FIGURE 3-33. Boxplots of subregion average daily maximum ozone concentrations (pphm) for each objectively determined ozone pattern (TESPATT; "." = missing).

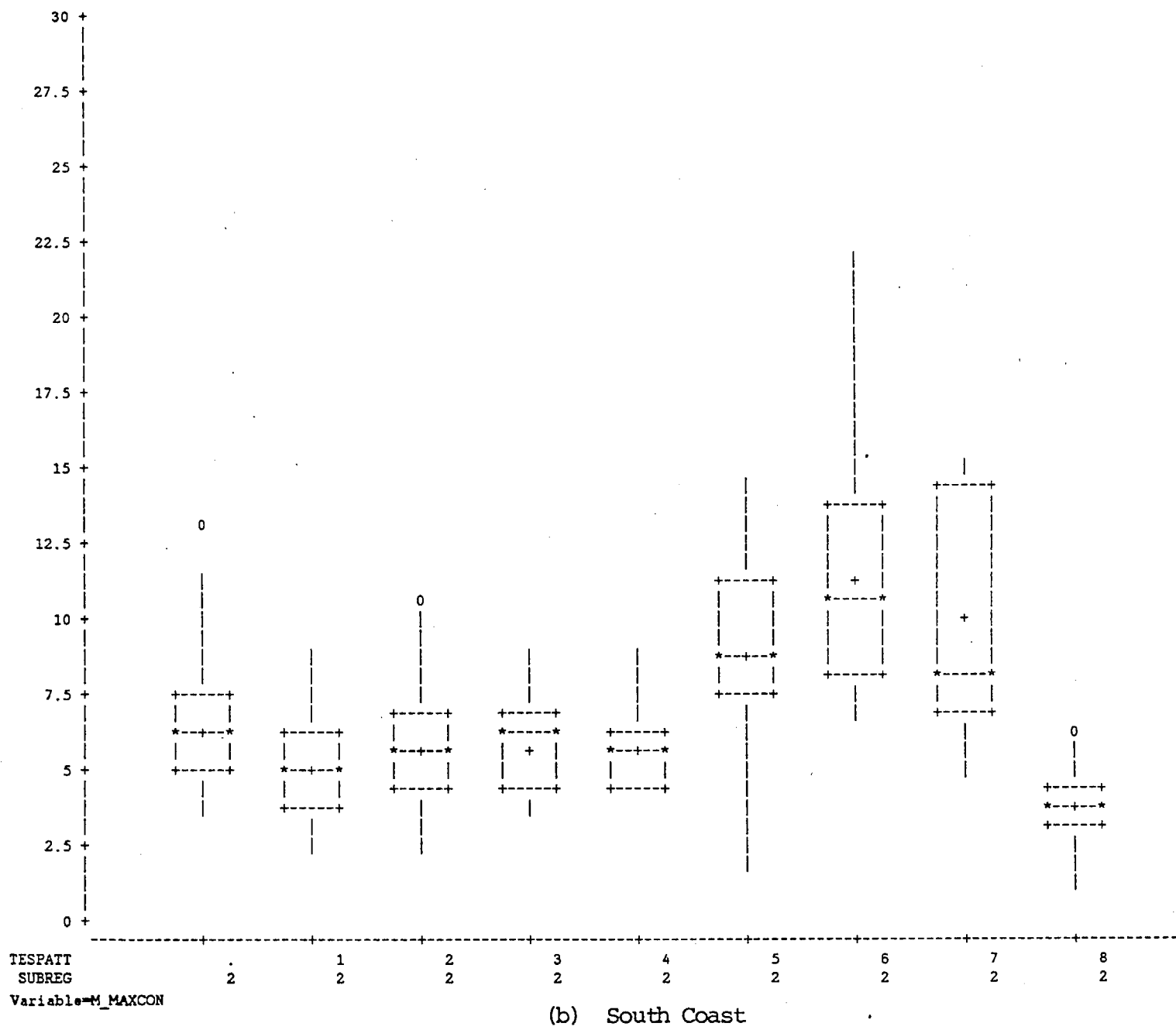


FIGURE 3-33. Continued.

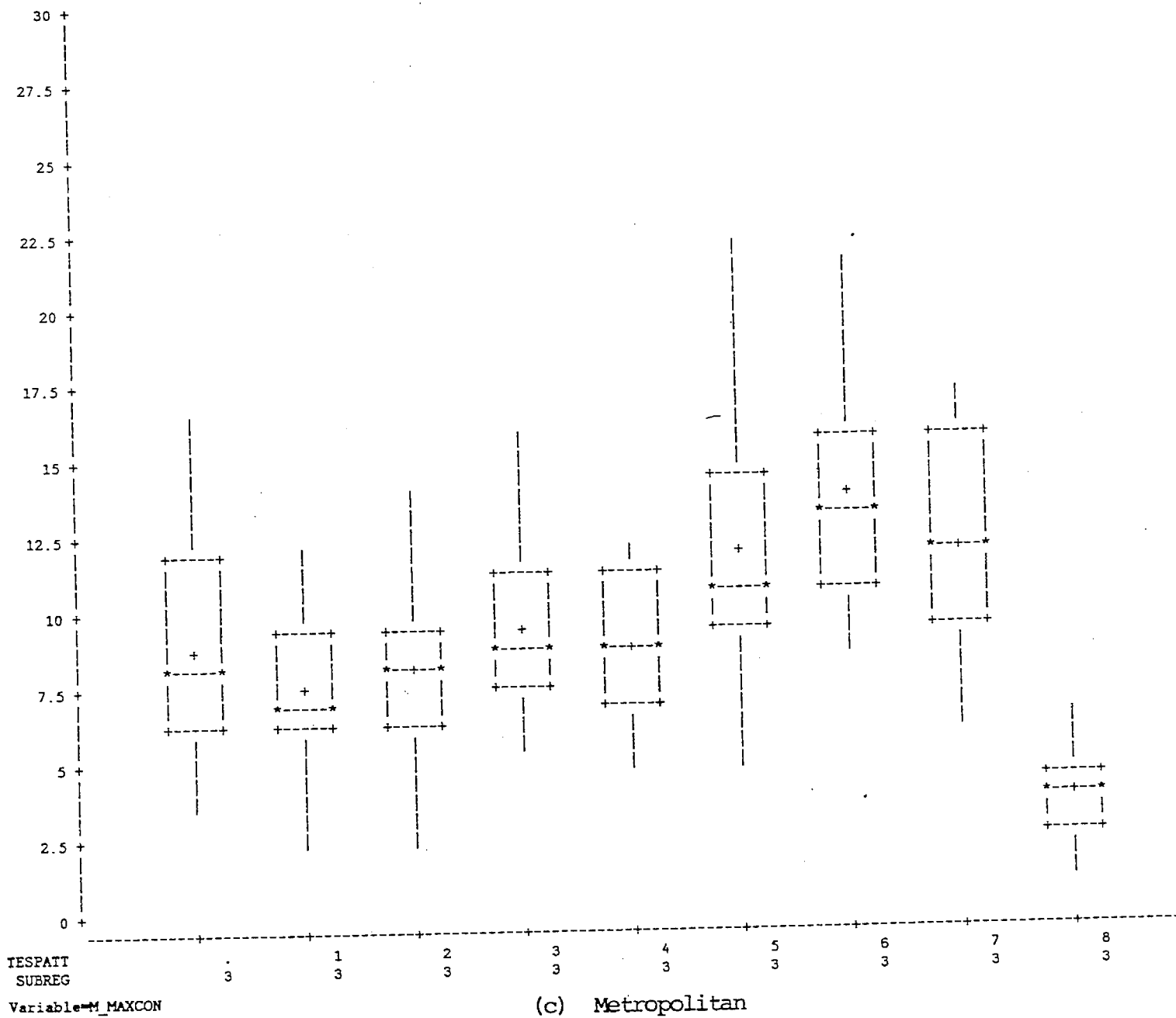


FIGURE 3-33. Continued.

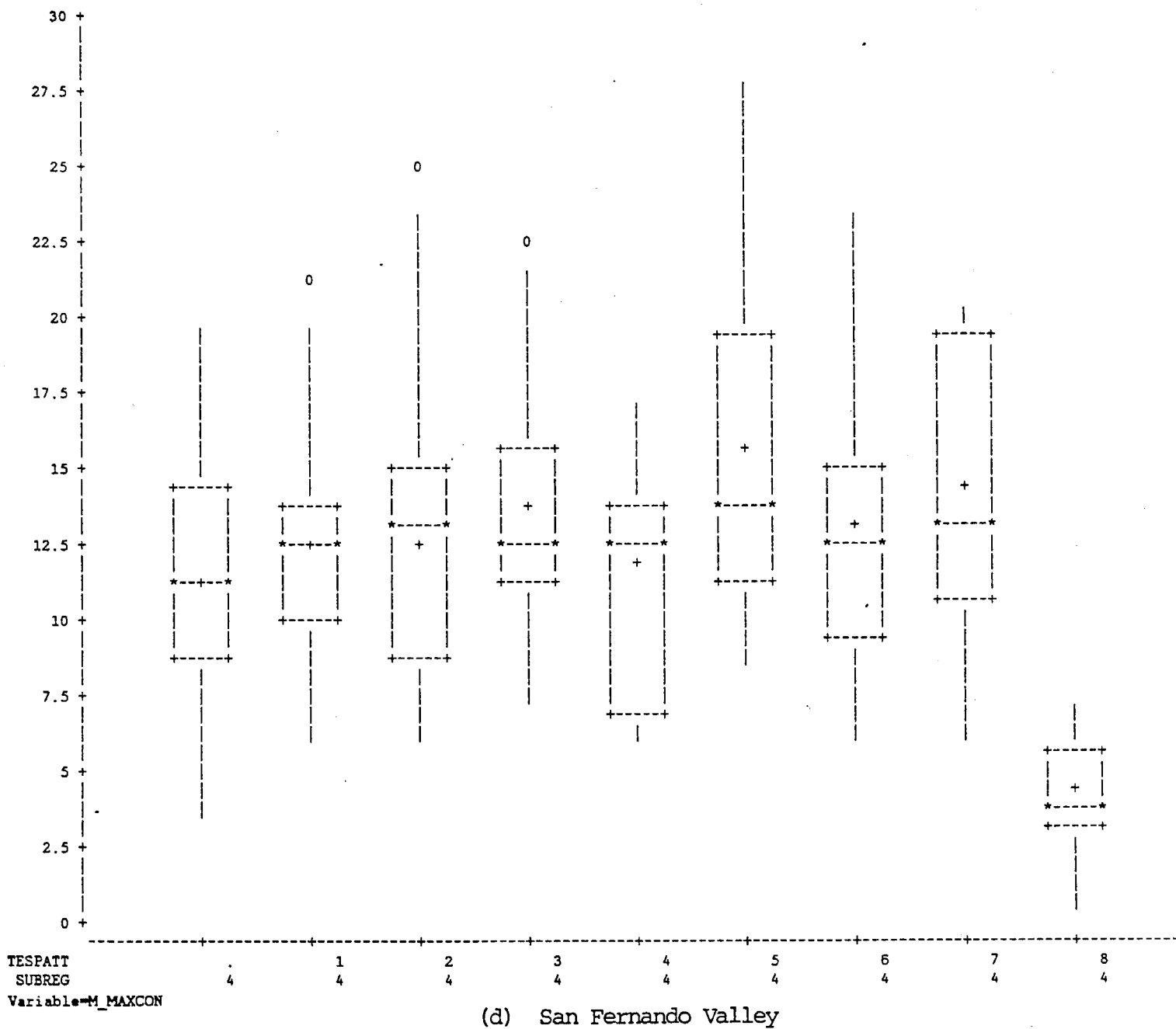


FIGURE 3-33. Continued.

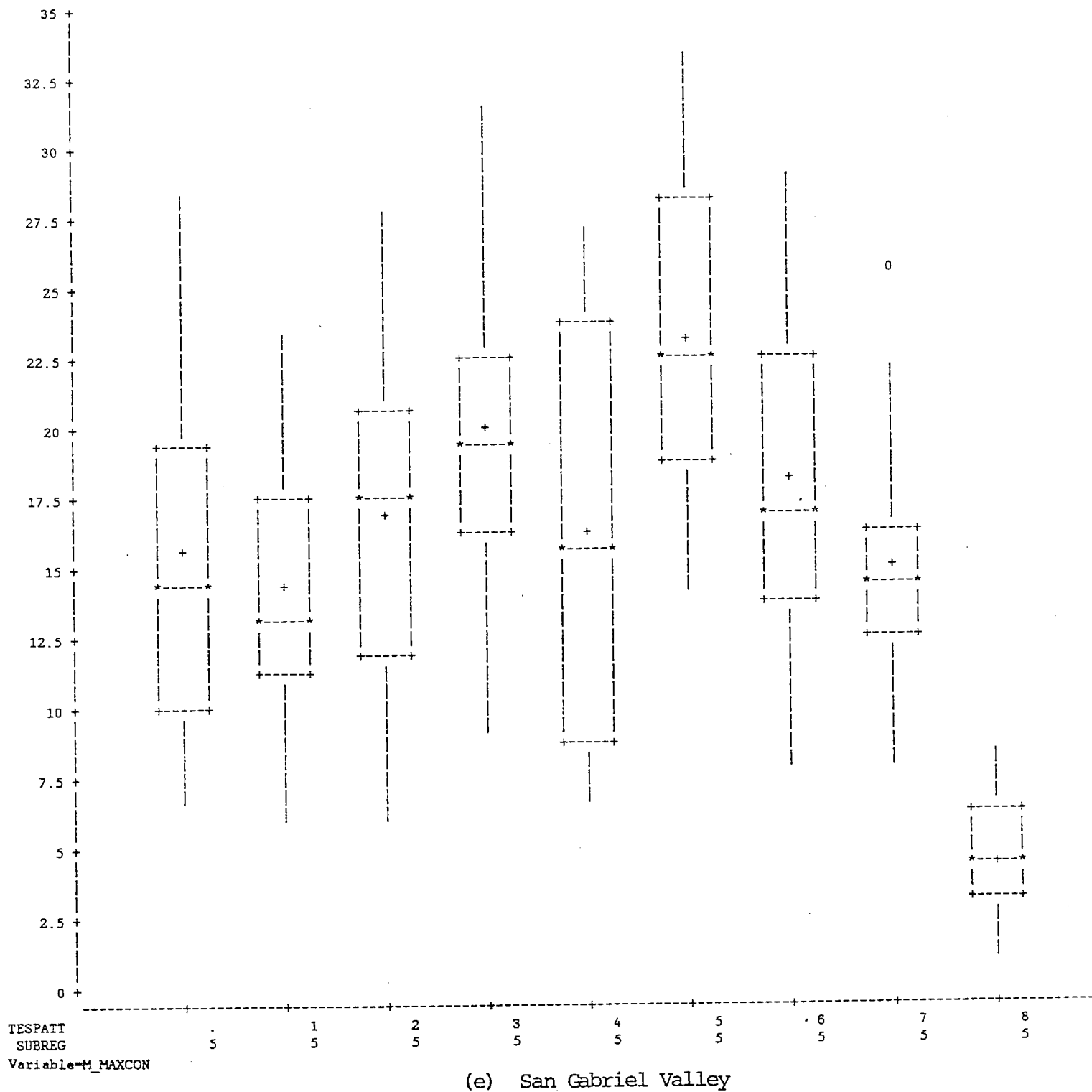


FIGURE 3-33. Continued.

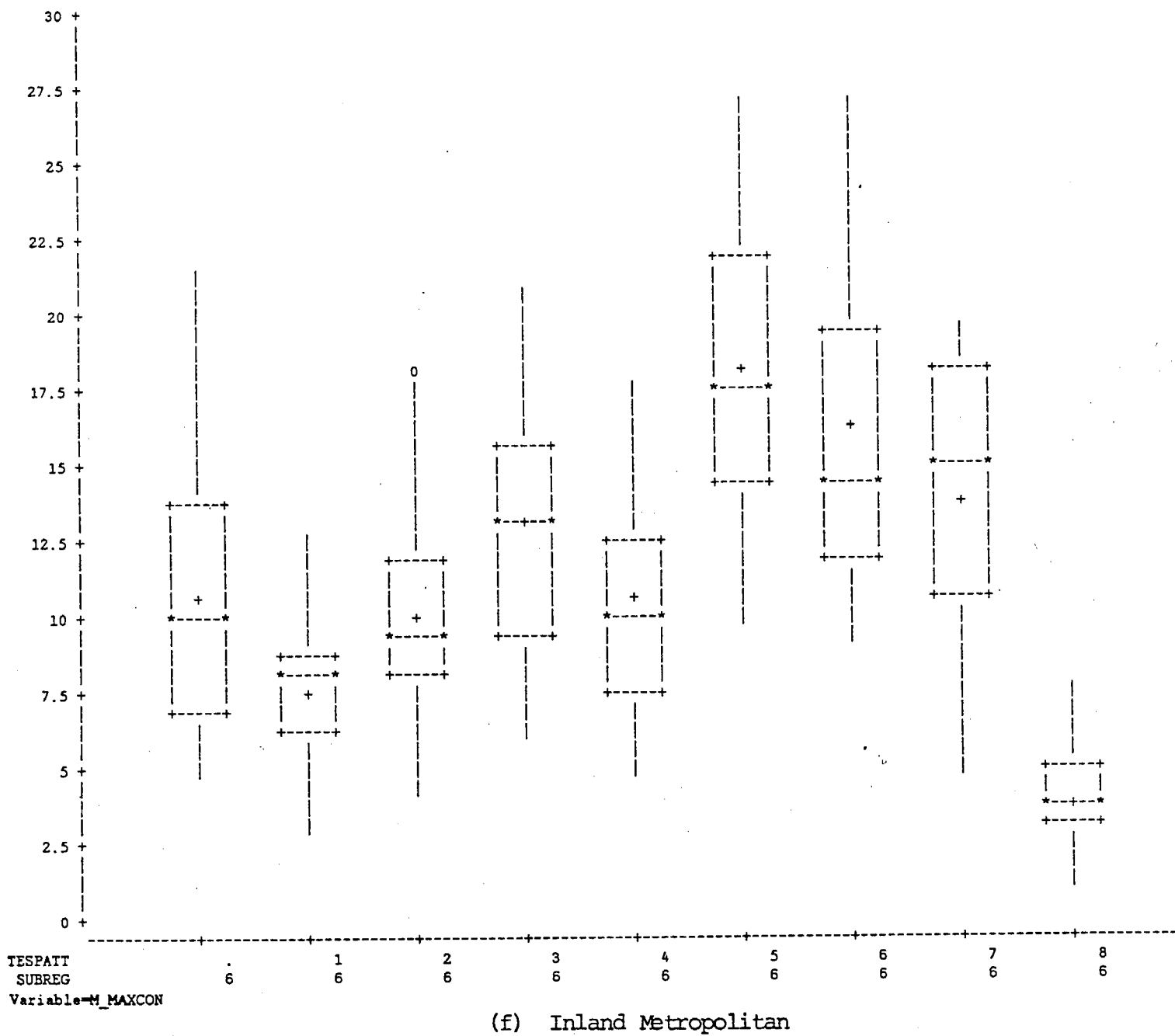


FIGURE 3-33. Continued.

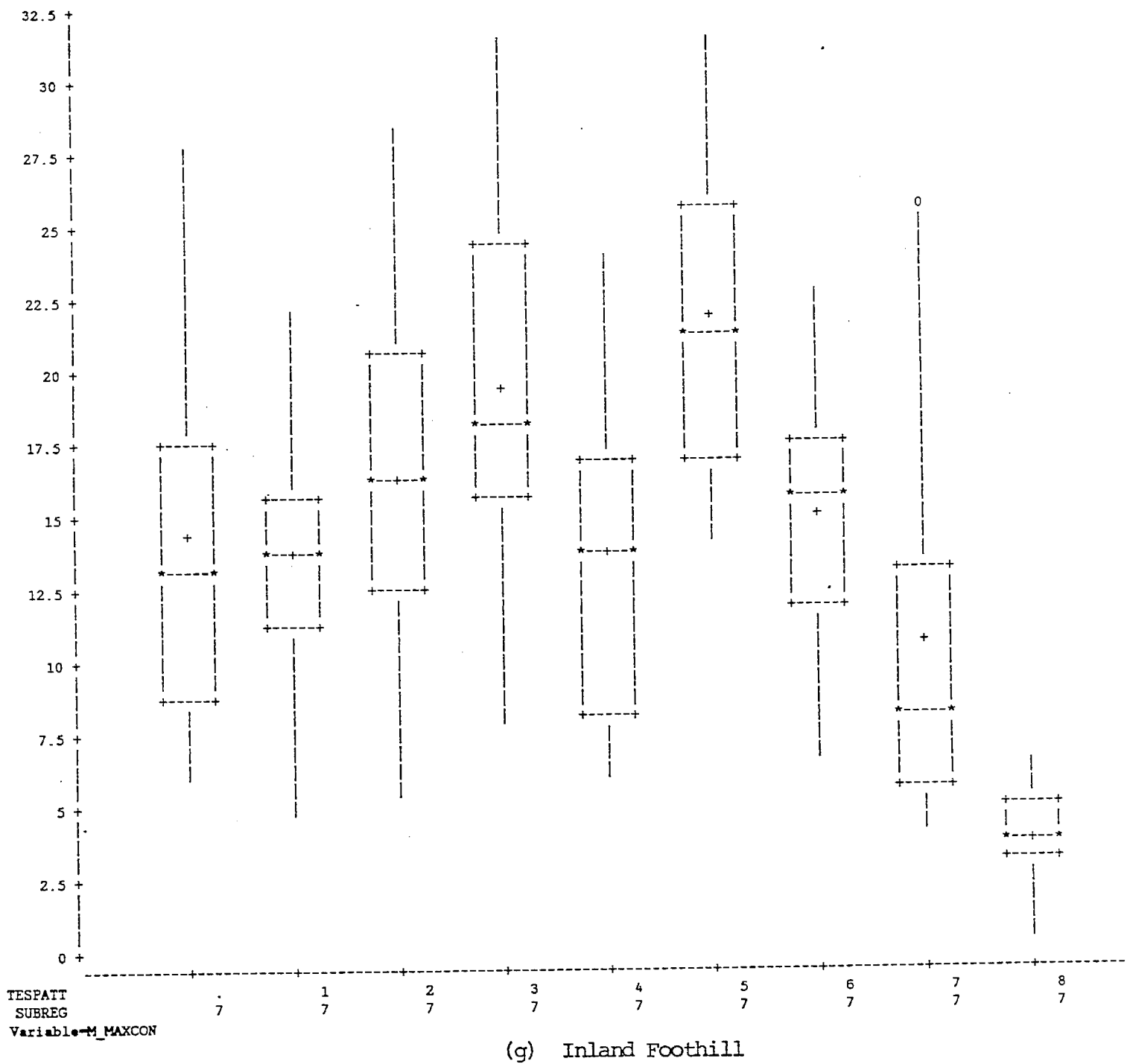


FIGURE 3-33. Continued.

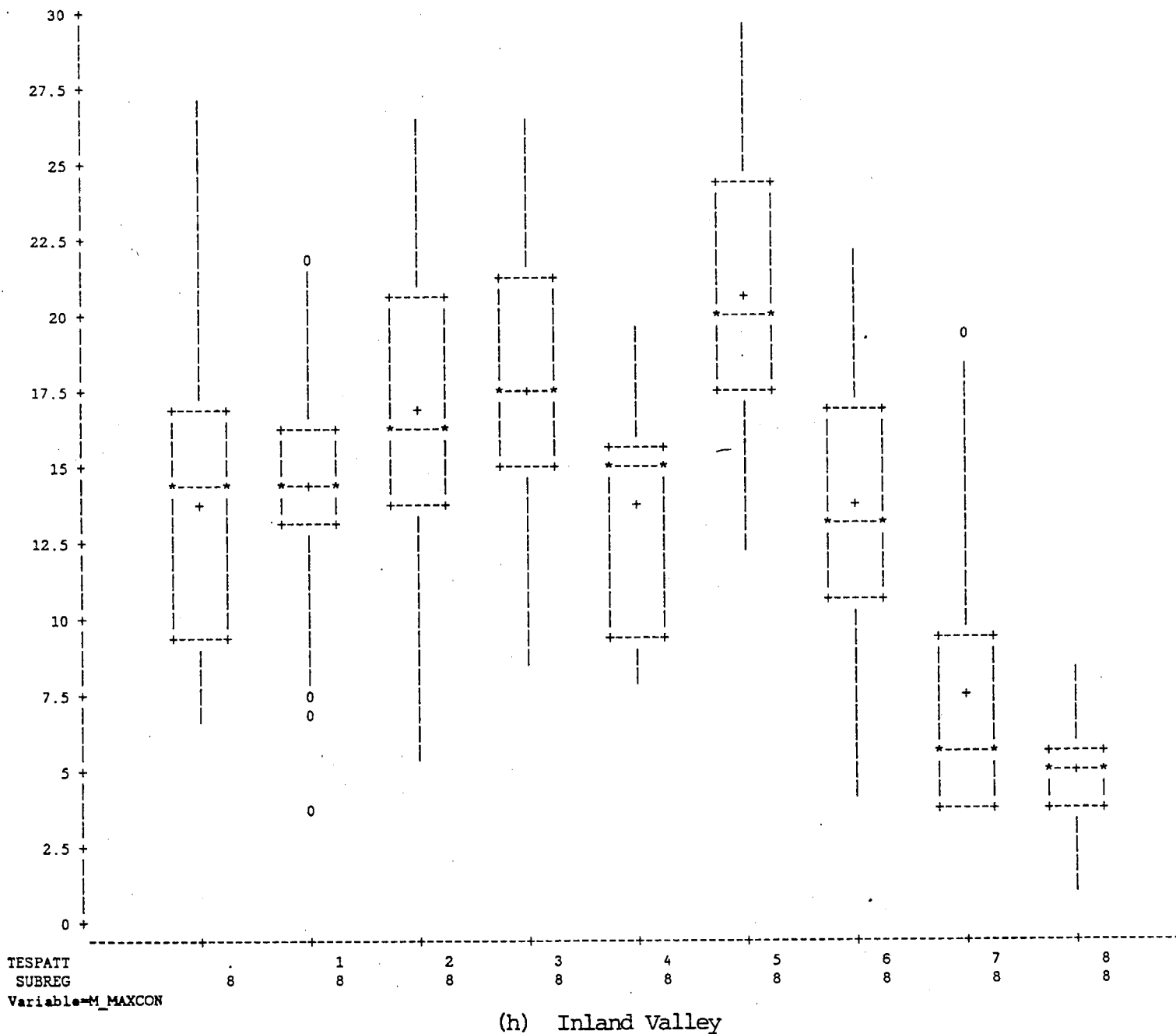


FIGURE 3-33. Continued.

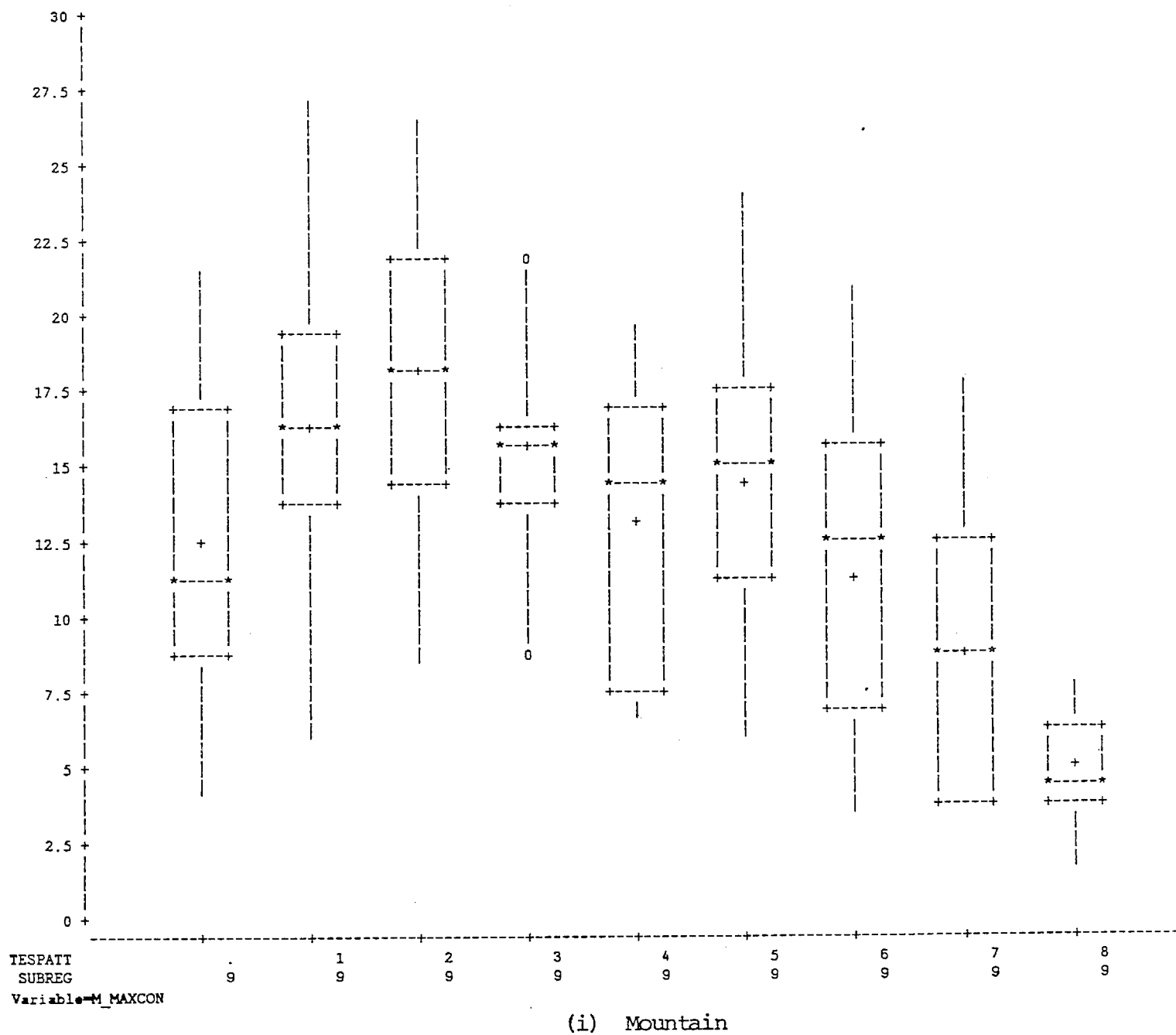
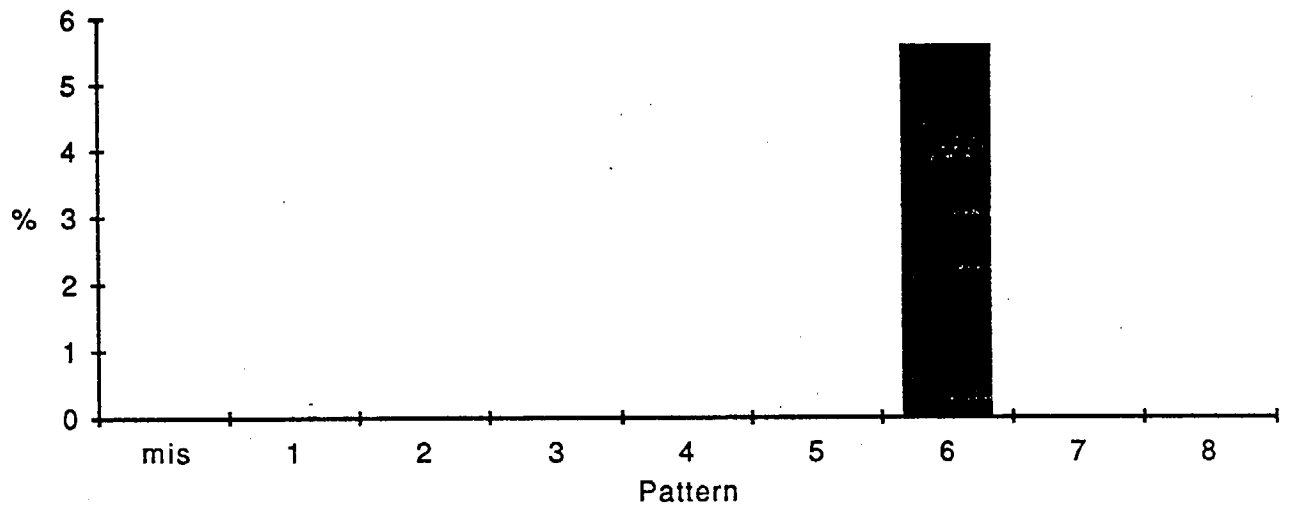
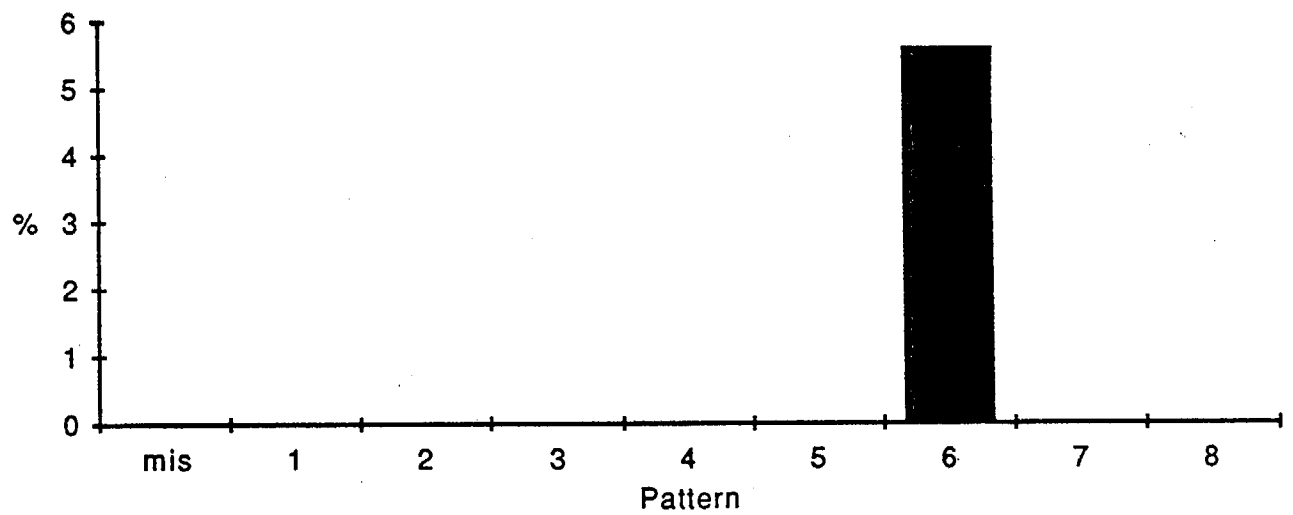


FIGURE 3-33. Concluded.

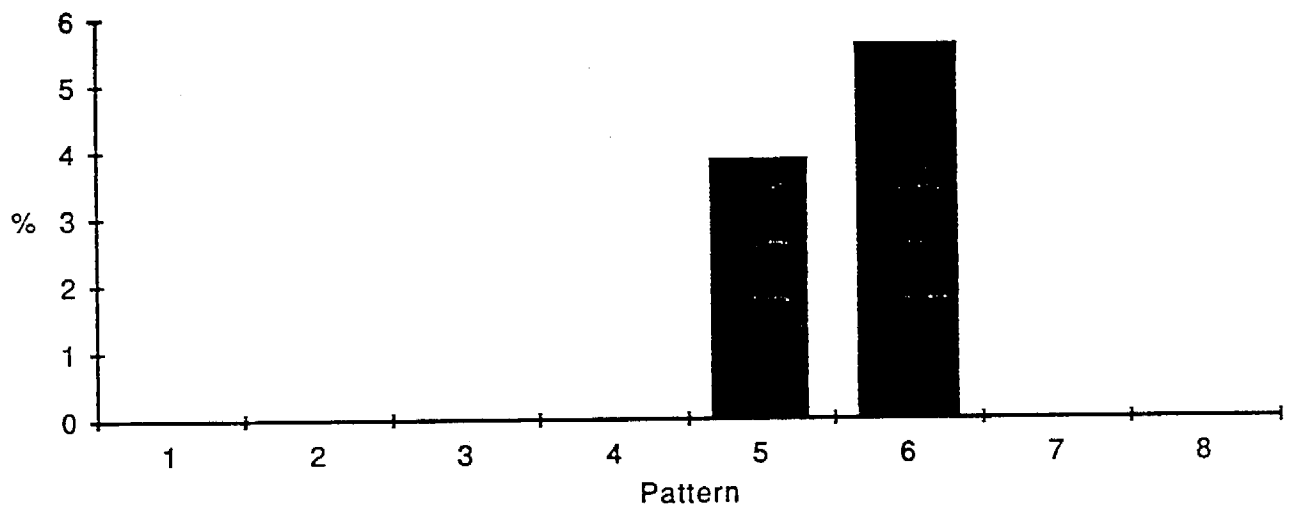


(a) North Coast

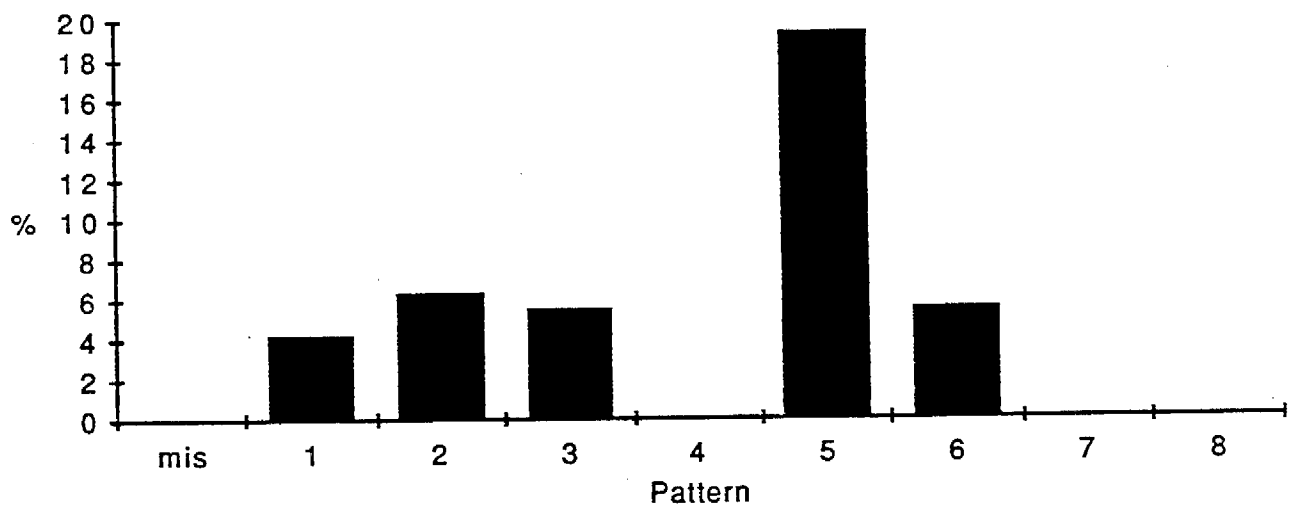


(b) South Coast

FIGURE 3-34. Percent of days in each ozone pattern on which subregion average daily maximum ozone concentration exceeds 20 ppm.

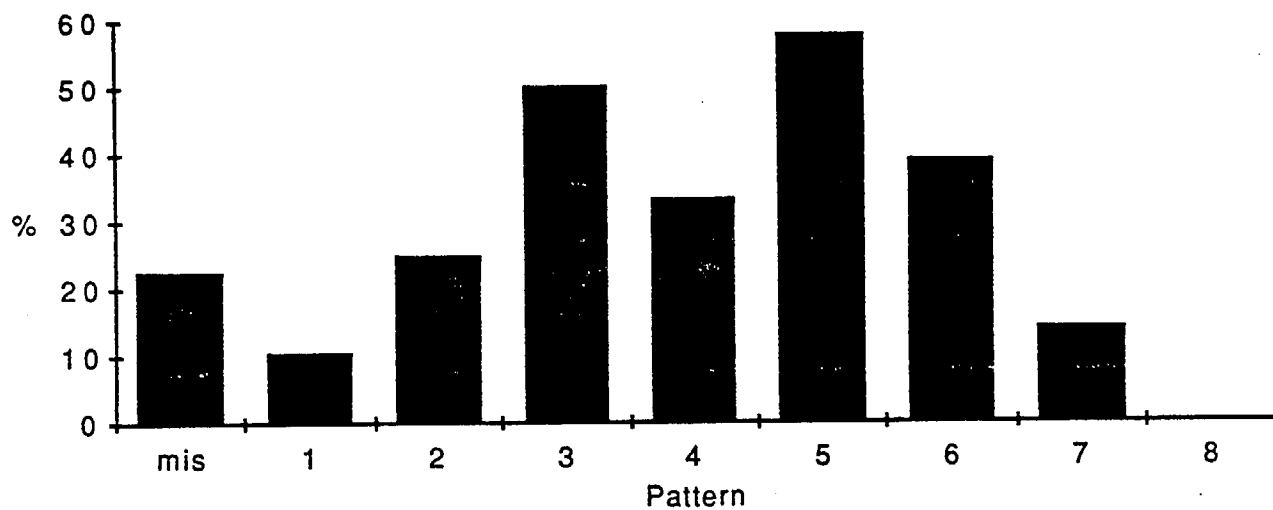


(c) Metropolitan

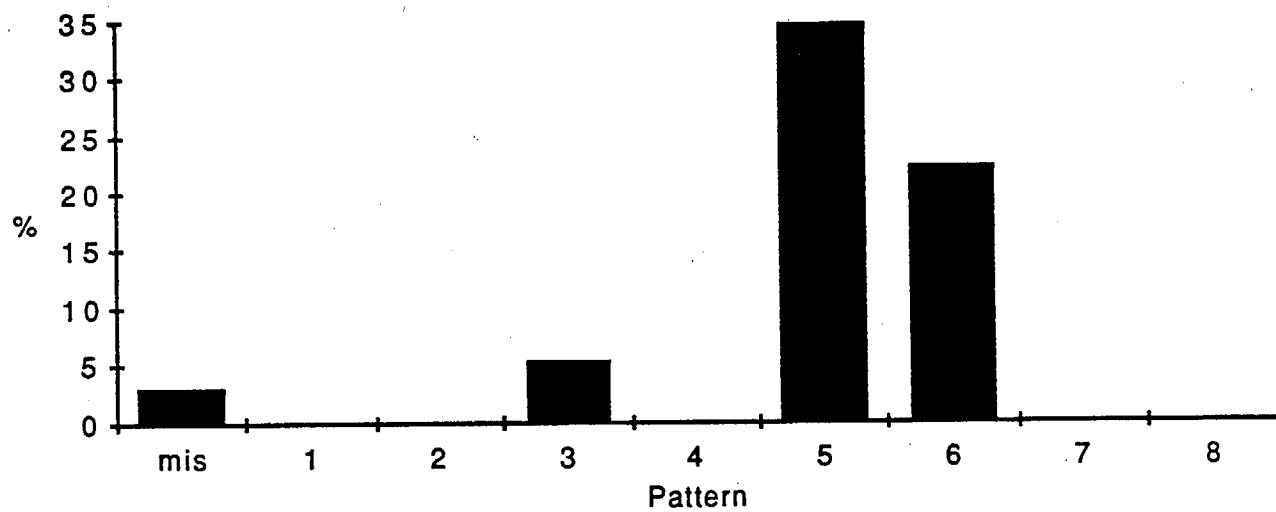


(d) San Fernando Valley

FIGURE 3-34. Continued.

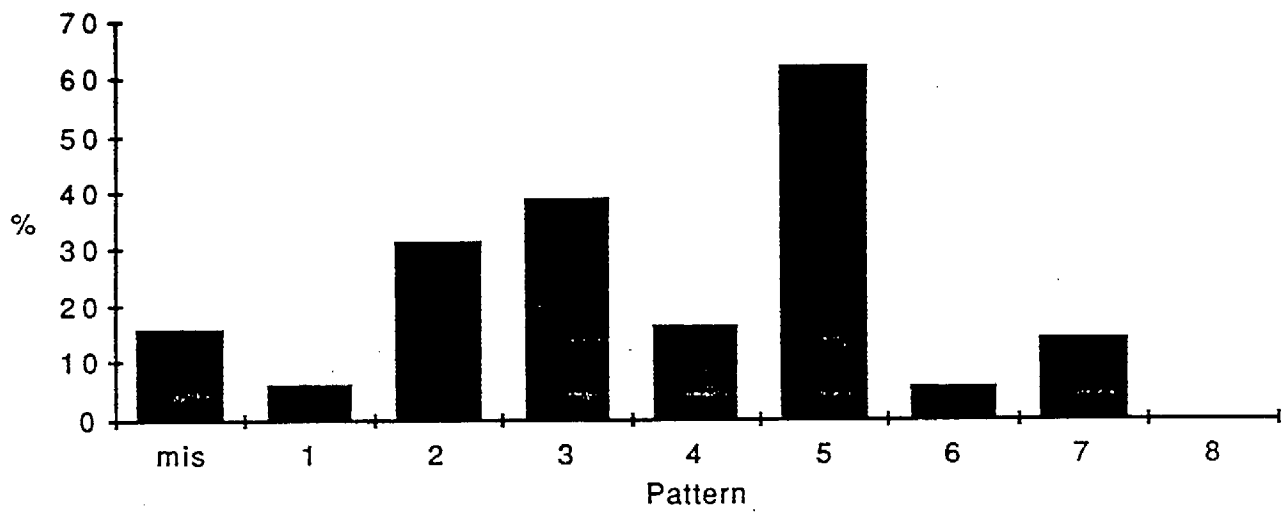


(e) San Gabriel Valley

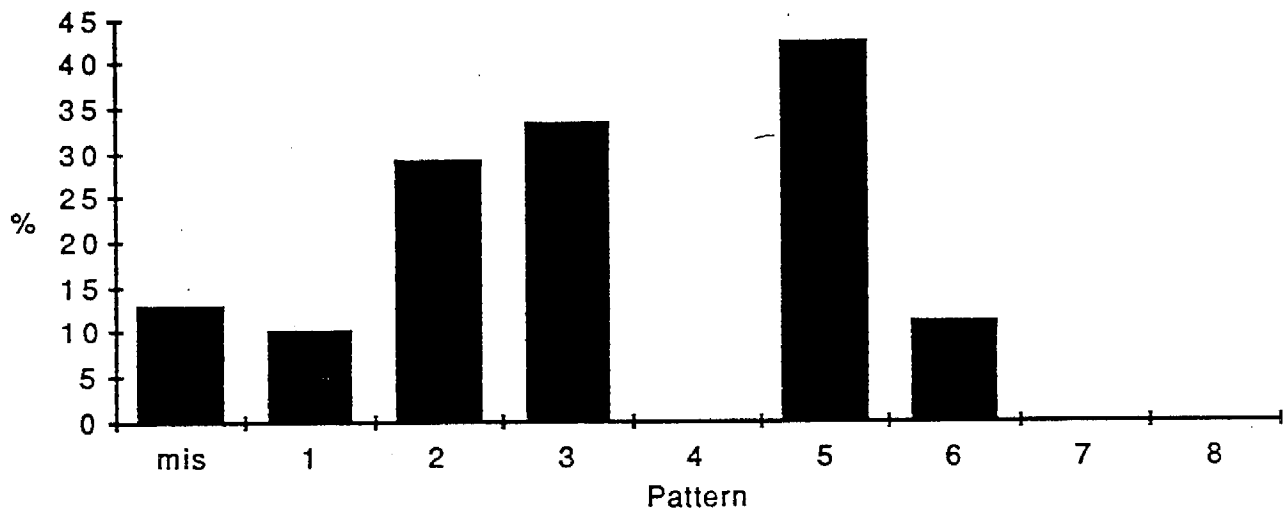


(f) Inland Metropolitan

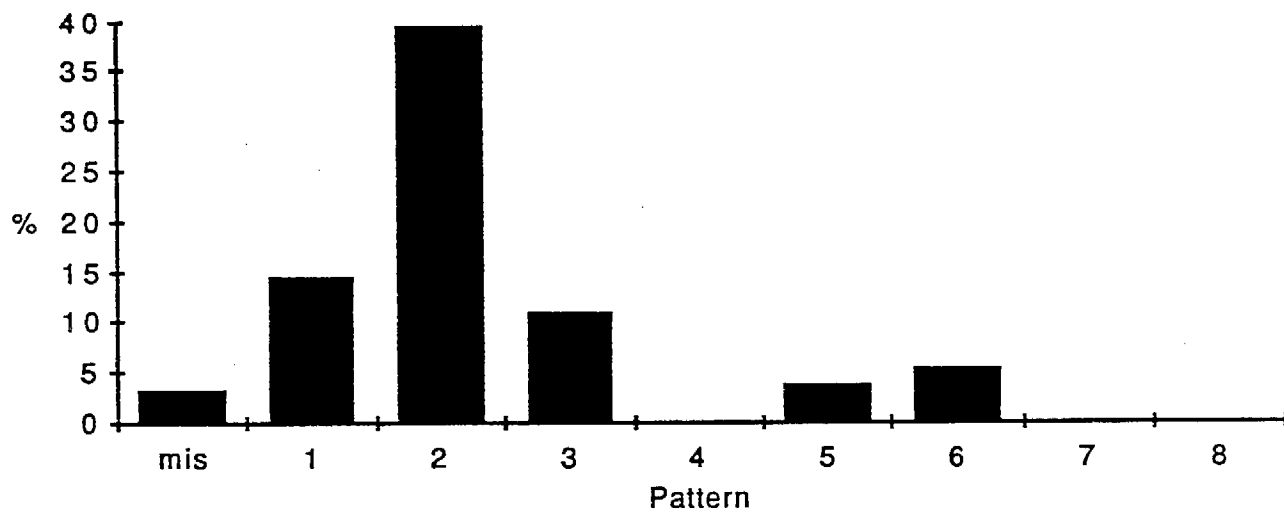
FIGURE 3-34. Continued.



(g) Inland Foothill

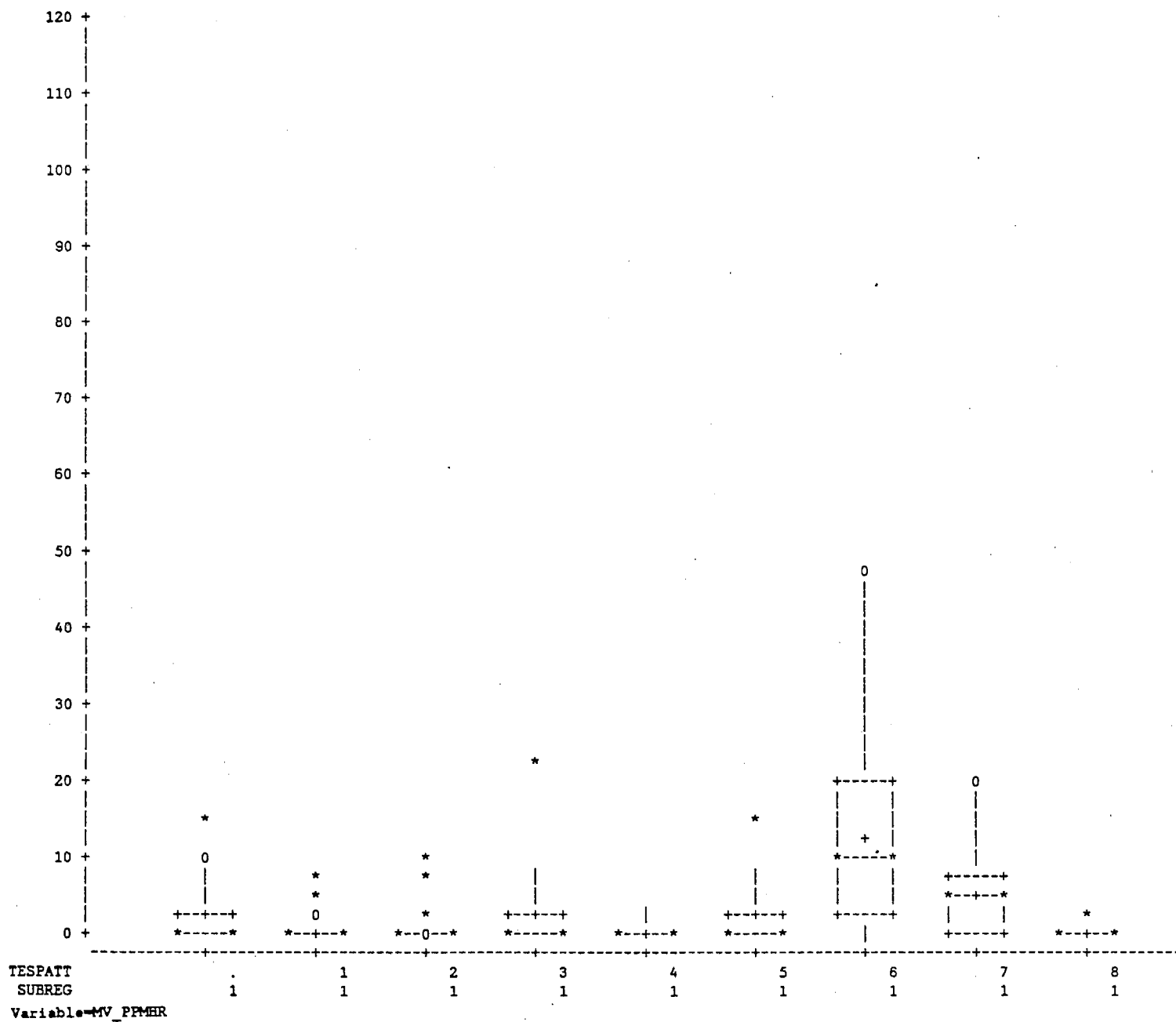


(h) Inland Valley



(i) Mountain

FIGURE 3-34. Concluded.



(a) North Coast

FIGURE 3-35. Boxplots of pphm-hours per day above 9 pphm for each objectively determined ozone pattern (TESPATT; "." = missing).

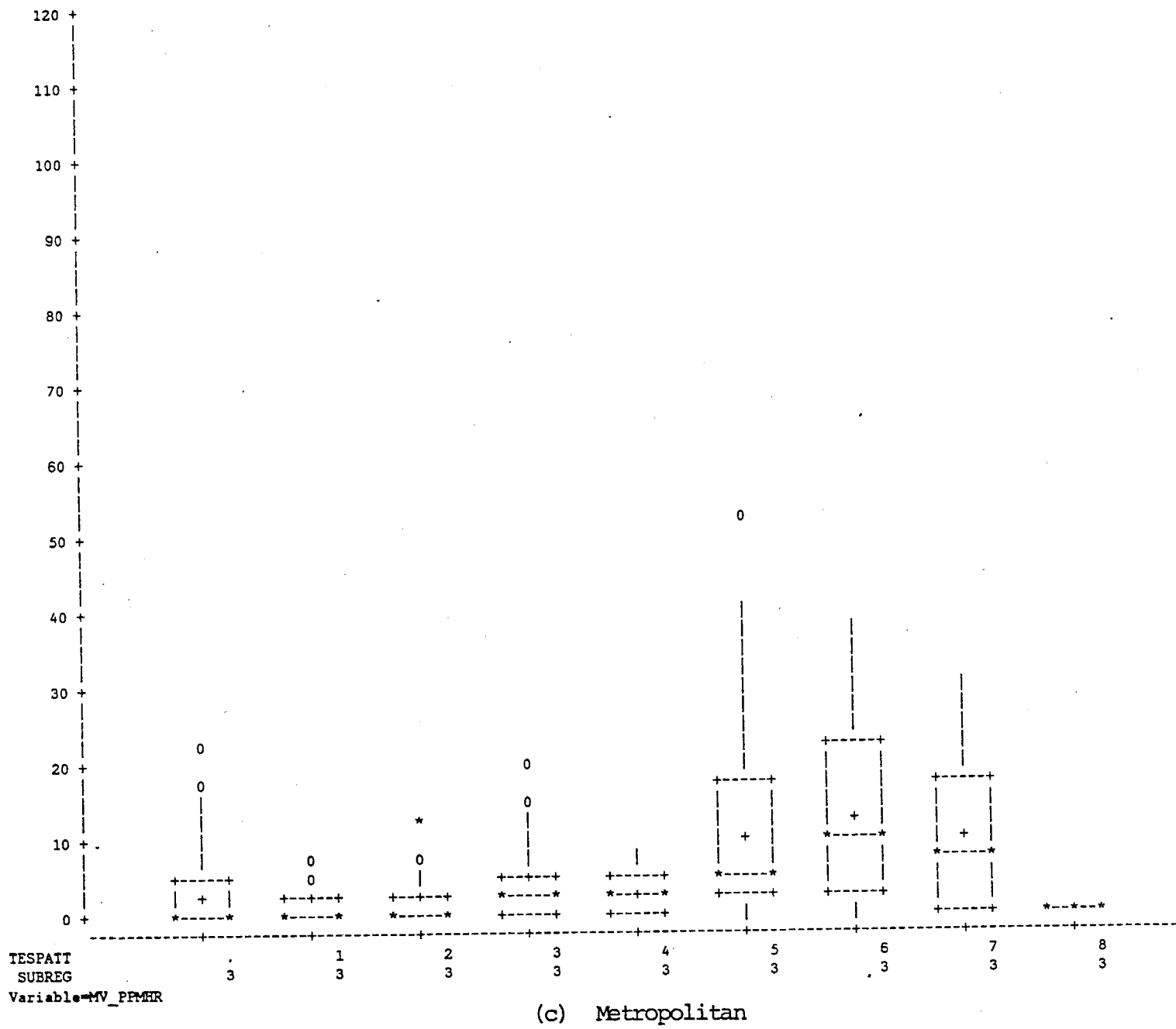


FIGURE 3-35. Continued.

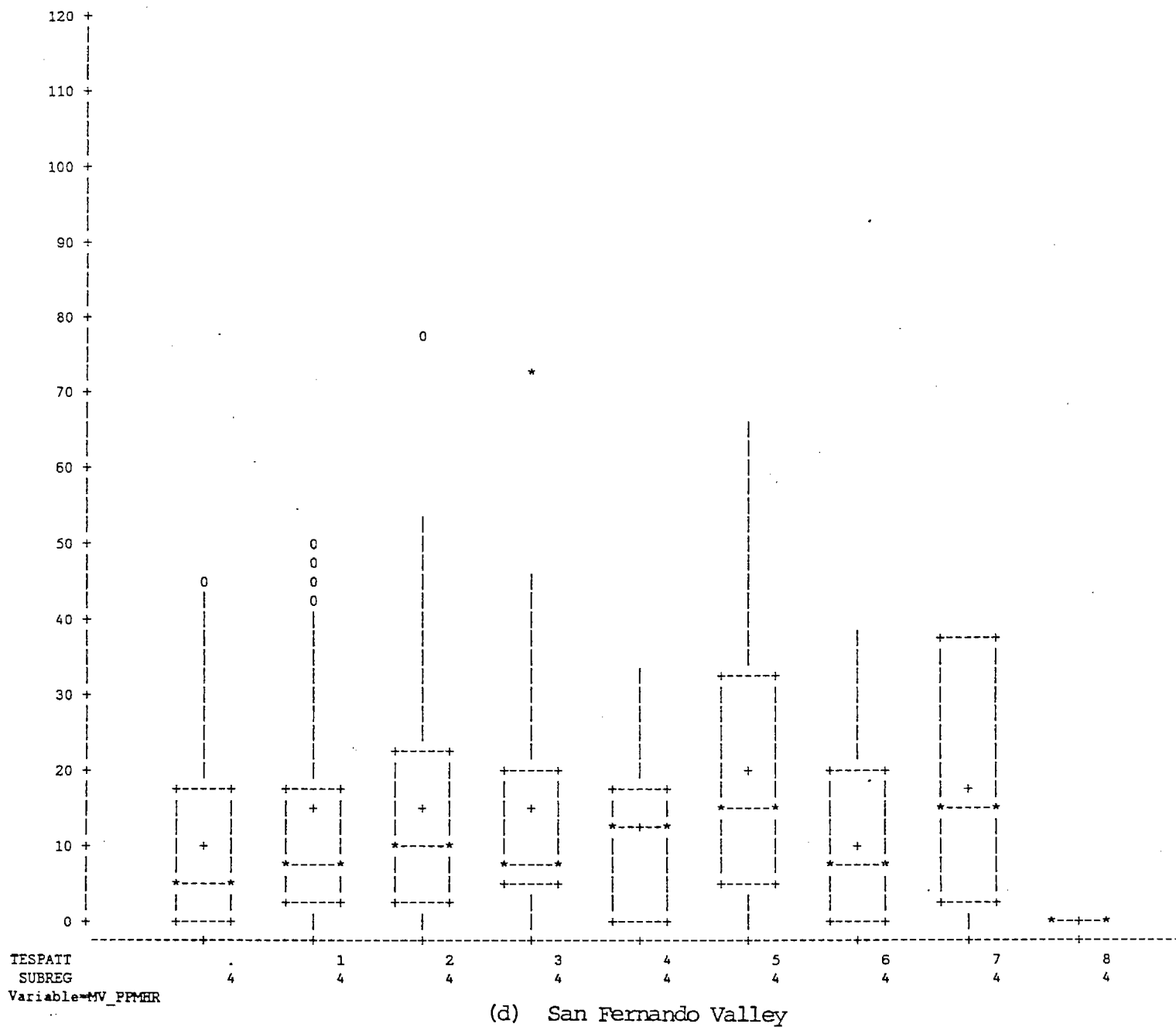


FIGURE 3-35. Continued.

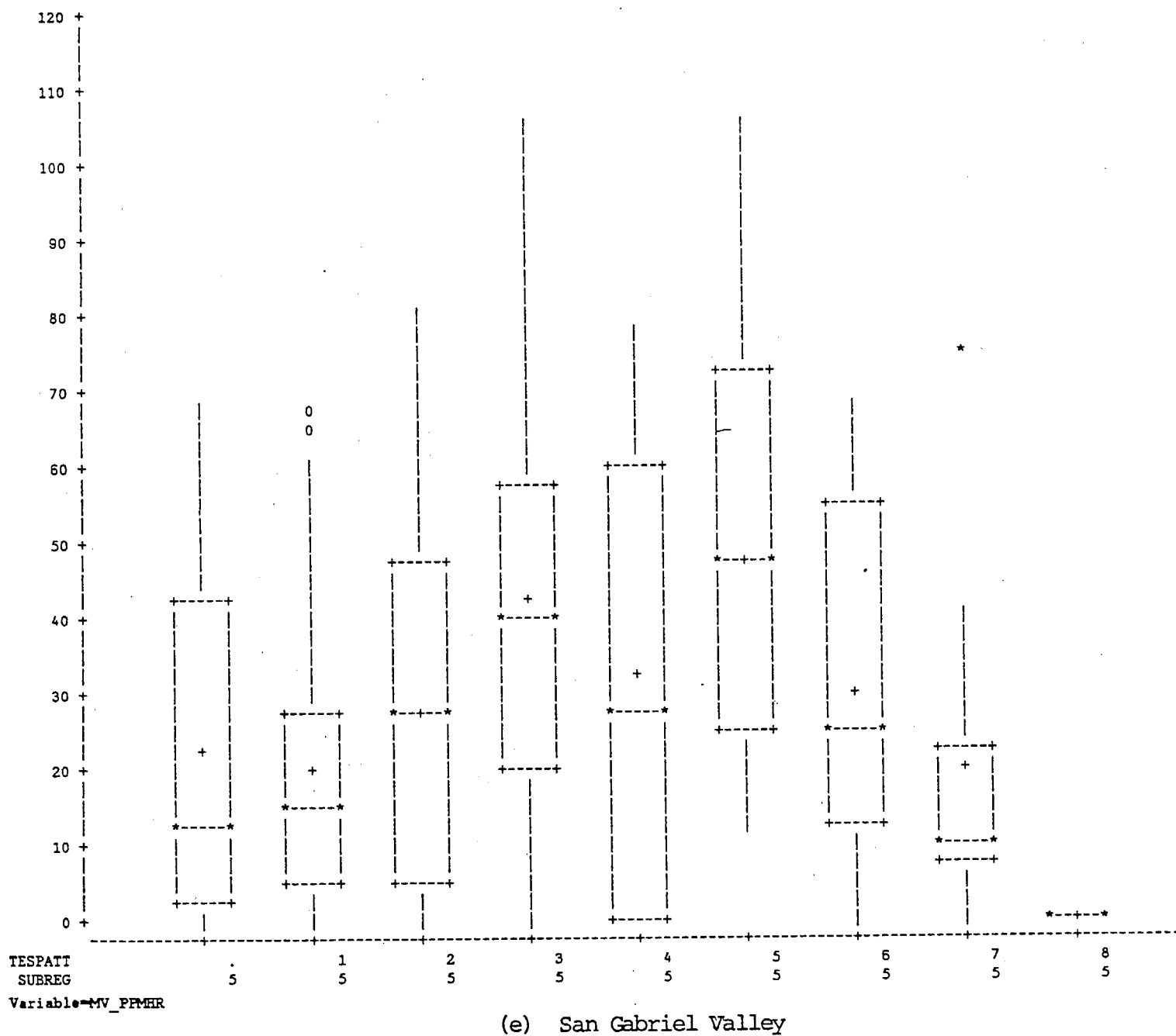


FIGURE 3-35. Continued.

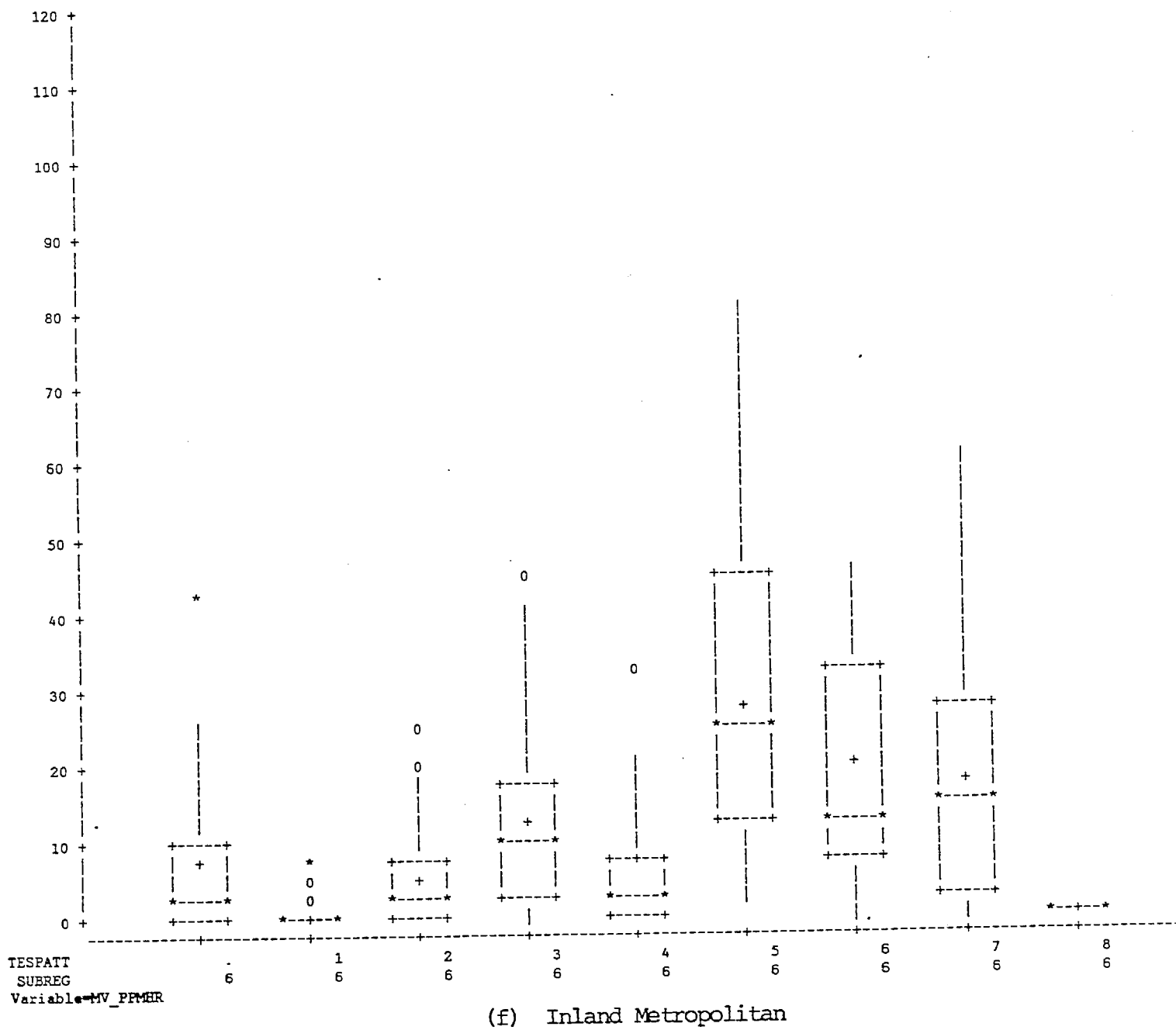


FIGURE 3-35. Continued.

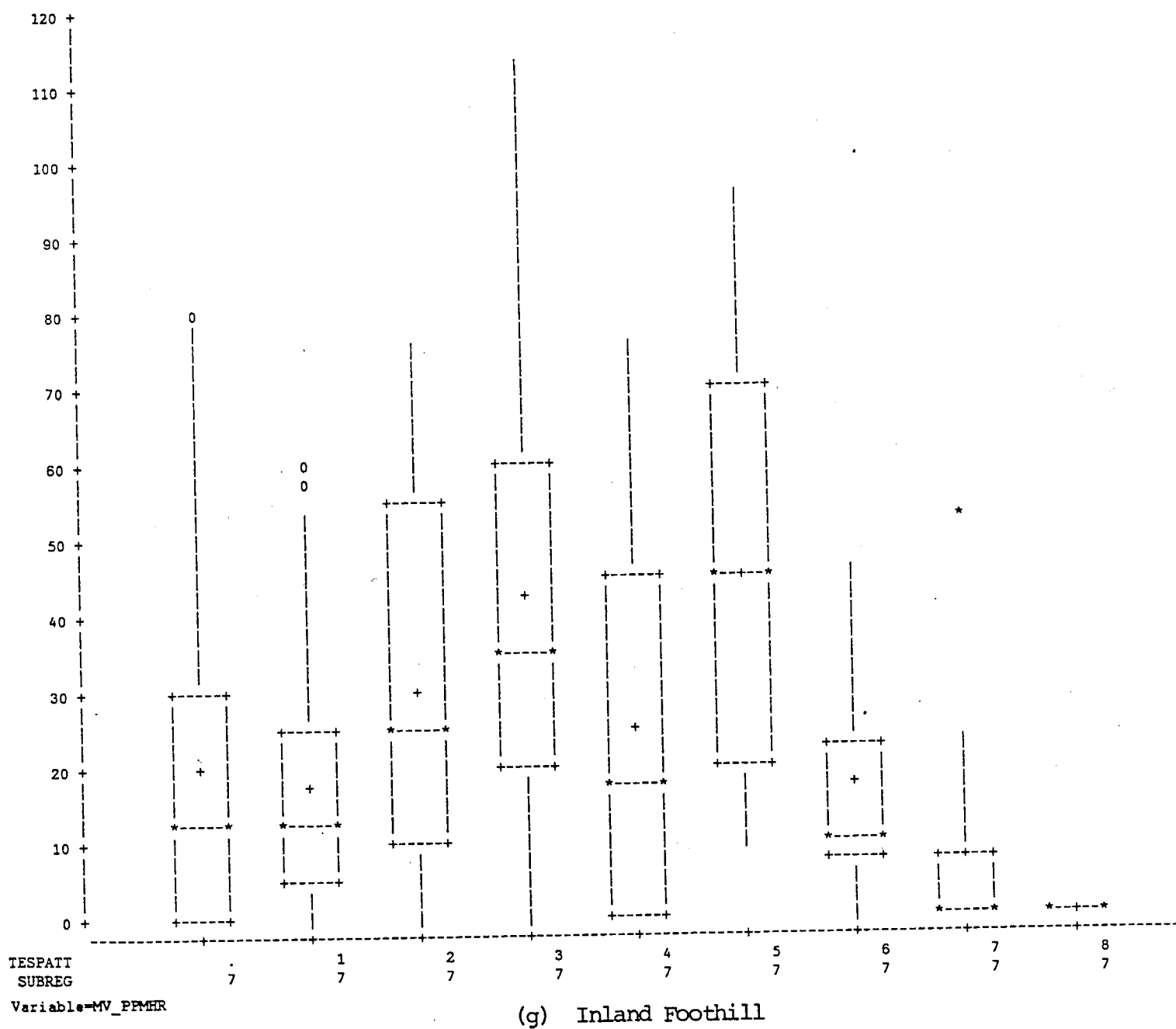


FIGURE 3-35. Continued.

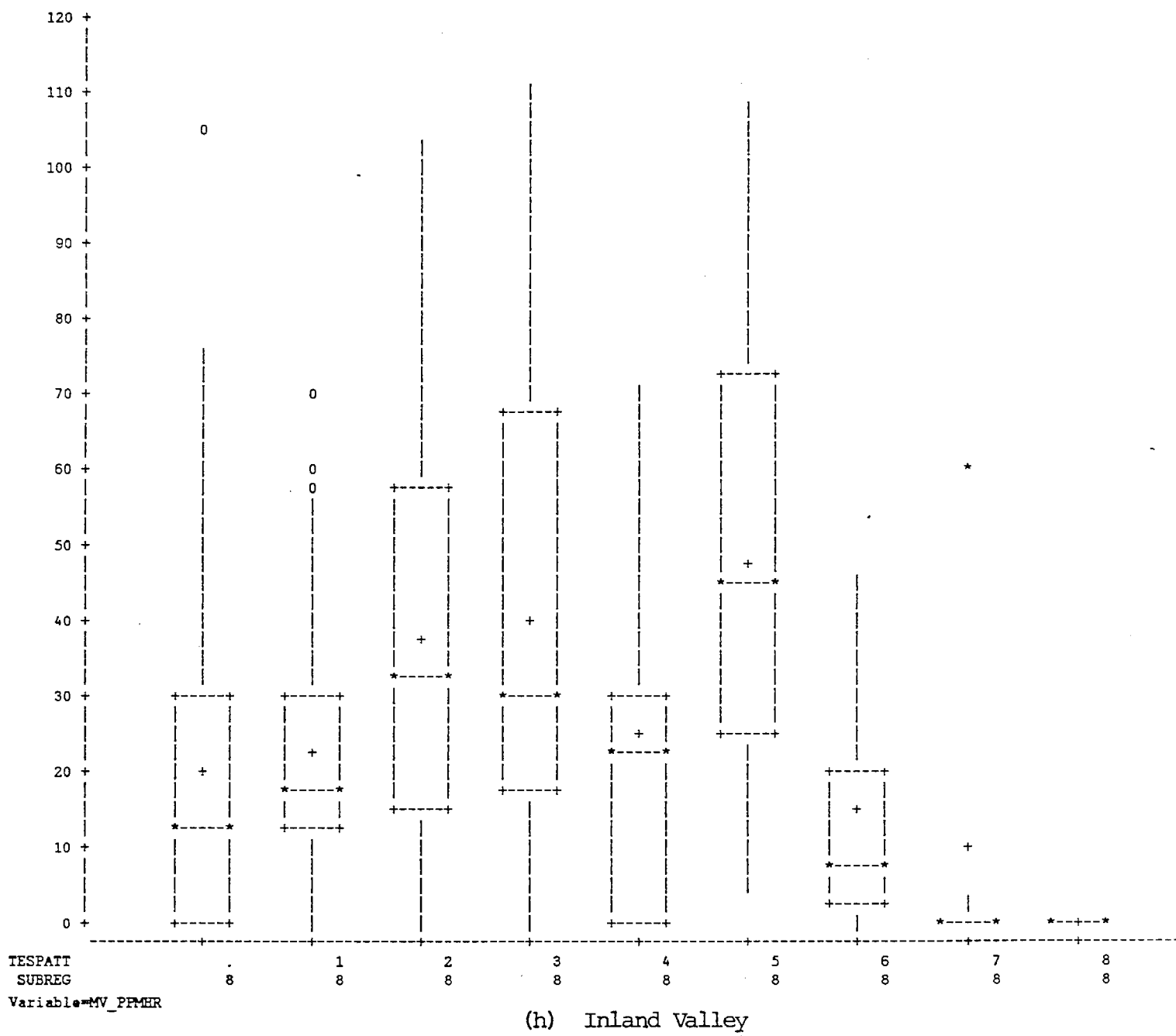


FIGURE 3-35. Continued.

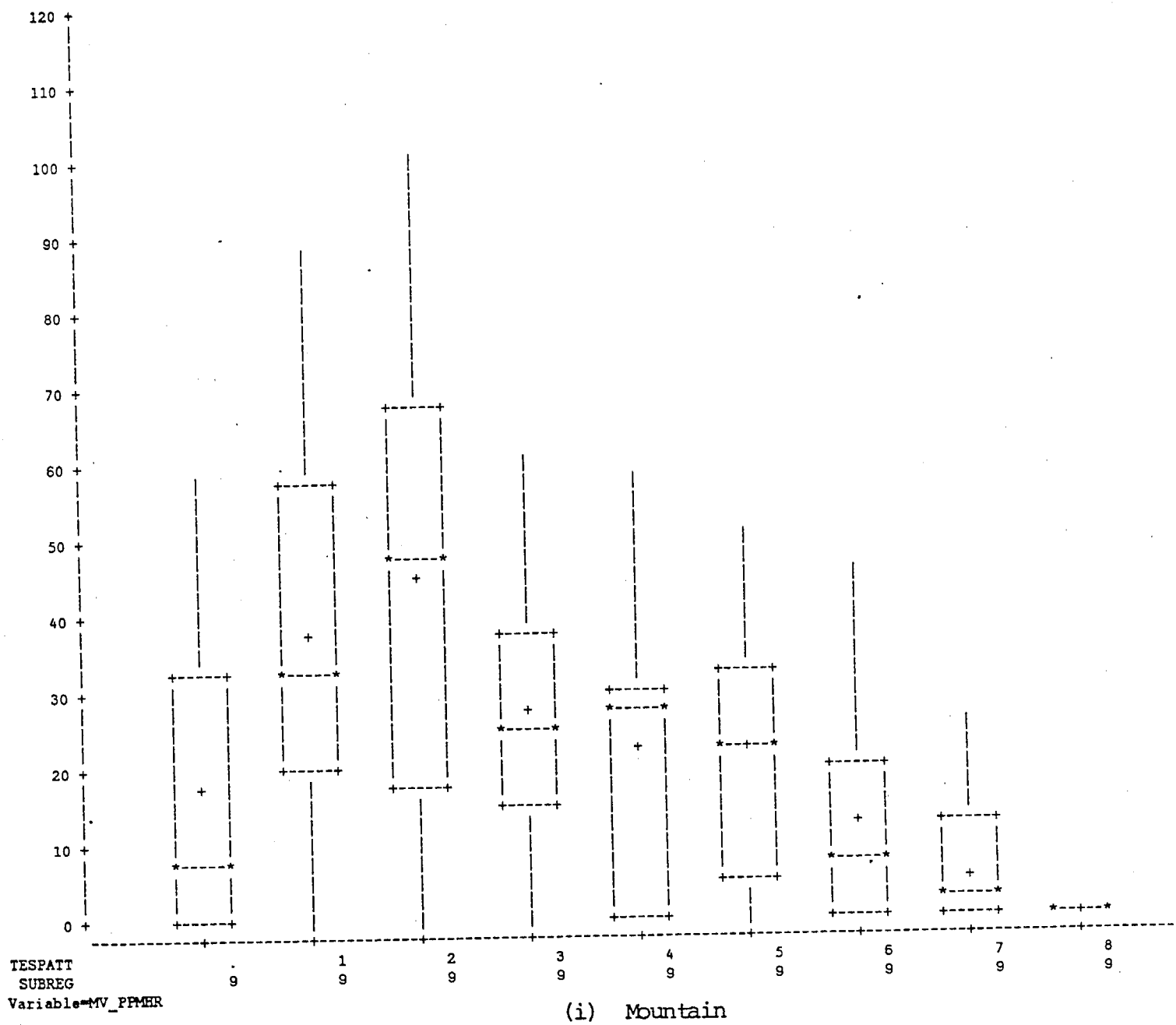


FIGURE 3-35. Concluded.

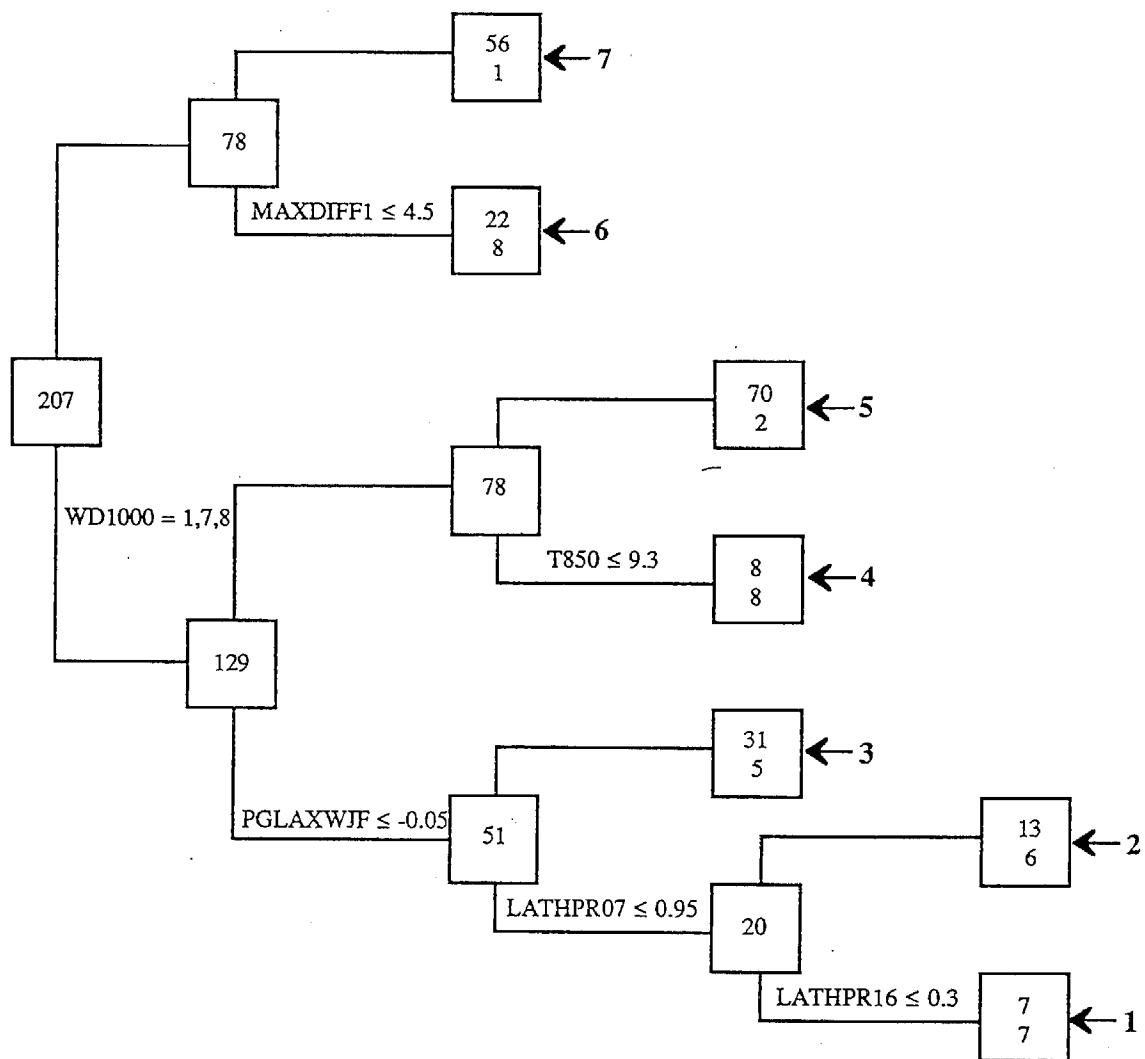


FIGURE 3-36. Classification tree illustrating relationship between objectively determined ozone patterns and meteorological variables (defined in Table 3-9). Numbers in boxes indicate number of days in each node. Bold numbers to right identify terminal nodes. Lower numbers in each terminal node indicate predominant source history category.